GENERAL LIBRARY
NOV 2 A 1918
UNIV. OF MICH.

Journal of Experimental Psychology

EDITED BY

JOHN B. WATSON, Johns Hopkins University
HOWARD C. WARREN, Princeton University (Review)
JAMES R. ANGELL, University of Chicago (Monographs)
SHEPHERD I. FRANZ, Govt. Hosp. for Insane (Index)
MADISON BENTLEY, University of Illinois (Index)

ADVISORY EDITORS

H. A. CARR, University of Colorado; RAYMOND DODGE, Wesleyan University; KNIGHT DUNLAP, The Johns Hopkins University; H. S. LANGFELD, Harvard University; E. L. THORNDIKE, Columbia University.

CONTENTS

Unidextrality and Mirror-Reading: Jos. E. Downey and Edwin B. Slosson, 393.

The Differential Spatial Limen for Finger-Span: Howard S. Langfeld, 407.

A New Olfactometric Technique and Some Results: Harry Kirke Wolfe and Benjamin Kilpatrick, 417.

The Memory Values of Mixed Sings of Alternating Series: Henry P. Angell, 435.

Children's Sense of Harmonies in Color and Tones: J. F. Dashiell, 459.

The Relation between Learning and Retention and Accuracy in the Learning Curve: A. C. Ellingson, 476.

PSYCHOLOGICAL REVIEW COMPANY
PRINCETON, N. J.

Journal of
Experimental Psychology

VOL. II, NO. 6. DECEMBER, 1917

UNIDEXTRALITY AND MIRROR-READING

BY JUNE E. DOWNEY AND EDWIN B. PAYSON

University of Wyoming

The suggestion out of which the present investigation grew was the possible correlation of expertness in mirror-reading and degree of one-handedness. This possibility has been indicated as a tentative conclusion by the outcome of a previous investigation.[1] In many ways such a conclusion is highly stimulating, since it suggests relationship between what might seem to be a purely sensory capacity on the one hand, and a motor function on the other. The capacity of individuals for mirror-reading varies very greatly as shown by the previous investigation. This capacity correlates with ability to write mirror-script with the practiced hand and with the ability to read and write inverted script. A growth factor also appears to be involved, since mirror-reading shows an inverse correlation, although not a high one, with age.

That expertness in handling spatial relationships of the symmetrical sort may be a function of the specialized anatomical organization which issues in unidextrality[2] is easily comprehensible. It is, however, a conclusion of considerable moment, particularly when considered in the light of possible variations in degree of handedness. So far as mirror-writing is concerned we were anxious to obtain more reliable data. The greatest obstacle was the difficulty in getting a satis-

[1] Downey, J. E., On the Reading and Writing of Mirror-Script, PSYCHOL. REV., 1914, 21, 408–441.

[2] We have adopted the suggestions of Ernest Jones concerning terminology. See PSYCHOL. BULL., 1909, 6, 130–132.

factory index of unidextrality. We have approached the subject from several different angles. Our experiments fall into two divisions; those carried out in the spring of 1916, reported in Part I., and those carried on throughout the academic year of 1916–1917, reported in Part II.

Part I

1. *Mirror-Reading Test.*—In our experiments in the spring of 1916 we obtained a measure for ability in mirror-script reading by duplicating the method of attack reported in a previous paper.[1] We tested individually a new group of subjects, twenty-five in all, by timing the rate at which two sentences written in formal mirror-script could be read and then ranking the subjects in an order of merit (see Column 7, Table I.). The tests were each stopped at the end of eight minutes. Subjects who failed to read the sentence with such a time limit were ranked on the basis of the number of letters read correctly. The two mirror-reading tests gave by the formula

$$\rho = 1 - \frac{6\Sigma D^2}{\eta(\eta^2 - 1)}$$

a reliability coefficient of .74 (P. E. .064). This coefficient is slightly higher than that reported in the preceding paper.

This method of individual testing is somewhat tedious and later we dropped it in favor of the group method, which will be described in the second part of the paper.

For obtaining the degree of unixdetrality we employed the steadiness test, the strength of grip test, and the comparative speed and quality of handwriting of the two hands. We had planned also to use the tapping test but were prevented by circumstances from getting these records.

2. *Steadiness Test.*—As one means of estimating the relative control of the two hands we tried Whipple's test 13, using the instrument pictured on page 157 of his Manual (Part I., second edition), with stop-watch, dry battery, telegraph sounder, key, and adjustable chair. Instead of

[1] Downey, J. E., and Anderson, J. E., Form and Position in Handwriting Interpretation, *J. of Educ. Psychol.*, 1915, **6**, p. 292 f.

employing the graphic method for registration of contacts, the experimenter counted them. Preliminary practice was carried to a point where we felt reasonably confident of our accuracy in counting. With regard to details of procedure, such as proper holding of the arm, placing of pencil in the hole, and the like, we conformed to the standardized instructions given in Whipple.

Since we wished to compare the relative steadiness of the two hands, a hole was selected where the subject's contacts with the supposedly inferior hand were within the range of accurate counting. As small a hole as possible was used in order to demand effort of control on the reagent's part. After the hole had been selected by preliminary trials, the subject was given the 'Ready' signal and placed his pencil in the middle of the hole; the stop-watch was then clicked. Three seconds after the signal, the circuit was closed by pressing down the key and the contacts counted from that point on for fifteen seconds. At the end of fifteen seconds, the key was released and the subject told to rest. The superior hand was first tested, then the inferior hand. Four records were taken for each hand for every subject. To obtain the relative steadiness of the two hands, the sum of contacts with the inferior hand was divided by the sum of contacts with the superior. An average was obtained of three days' records and on the basis of these averages the subjects were ranked, the subject showing the greatest difference in control in favor of the dominant hand was placed first and the one showing least difference last.

In another series of tests, steadiness was determined by the length of time the pencil could be held in the chosen hole without contact. The record was made in seconds, beginning two seconds after the 'Ready' signal. The sum of seconds in which the inferior hand could be held in position without making contact was divided by the sum of seconds in which the superior hand could be so held. The order of merit obtained by this method correlated with that obtained by the first, + .67. Although the inaccuracy incident to counting of contacts was overcome by using the time-method, we

felt that, on the whole, it was less reliable than the former method because of the greater effect upon the record of a single contact. We have, accordingly, used the first steadiness test as a basis for comparison with other tests. Table I. (Column 3) gives the average record of the subjects in terms of the percentage of contacts made by the inferior hand when compared with the superior hand. The higher the percentage, the greater the difference in motor control between the two hands. One hundred per cent. would show equality in control and anything below this would indicate that the supposedly inferior hand was steadier than the dominant hand.

The utilization of the results of the steadiness test as a measure of the degree of handedness might, of course, be questioned; any one test would seem inadequate. Still more dubious might seem the arrangement on this basis of the individuals in an order of merit. Let us, however, study the records in the light of what we know of our subjects from other sources.

Subjects Wa, As and Dv are definitely left-handed, using the left hand for writing as well as for all other skilful operations. Subject Aw described himself as naturally left-handed although writing with the right hand because of outer compulsion. Our impression from conversation with many left-handed individuals leads us to believe that naturally left-handed persons resist very unequally the attempt to make them write with the conventional hand. In the same family, parental and pedagogical pressure may succeed with one child and fail with another; degree of handedness would appear operative as one factor determining degree of opposition. It seemed very possible that Aw was less left-handed than the other three left-handed subjects. A fifth subject, DJ, reported that she uses the left hand to a much greater degree than does the average person; in many situations she employs either hand with comfort.

Reference to Table I. shows that these reports of these five subjects were sustained by the outcome of the steadiness test. With one exception the other subjects appear to have

better motor control for the right hand, but in various degrees. Hi gives a better record for the left than for the right hand. Besides DJ, reagents McK, Ba and Pe show very slight superiority for either hand, as measured by motor control.

An interesting feature of our complete records concerns the matter of variability in the records from day to day. On one day each Pa, Pe, Ba, DJ and Hi gave a better record for the left hand. This aspect of the situation deserves very careful investigation, particularly in connection with its relation to the general condition of the subject. Our records are not extensive enough to justify a report.

3. *Strength of Grip.*—Strength of grip was measured by means of Smedley's dynamometer. Four trials with each hand, alternating the hands used, were given on each of two days, every effort being made to get the reagent to exert his maximal strength, with fair success, as shown by the slight variability. The percentage of strength of the inferior hand is compared with that of the superior hand. In getting this percentage we utilized the average of the eight trials. Table I. gives the results. Three of our four left-handed subjects have a stronger left than right hand; one, Wa, is stronger in the right hand. Of our right-handed group three, McG, DD and Sa show greater strength with the left hand.

Other investigators report inconsistencies in the outcome of the dynamometer experiment as an index of handedness. Mrs. Woolley from her observations on working children concludes that "superior strength of the left hand, taken alone, is no indication of left-handedness." And again, "steadiness constitutes a far better indication of left-handedness than strength, though it is by no means infallible." [1]

4. *Speed and Quality of Writing, Unpracticed Hand.*—In another series of tests we attempted a comparison of the relative skill of the group with regard to the two hands by testing writing-speed for both hands and finding the percentage number of letters written by the inferior hand when

[1] Woolley, H. T., and Fisher, C. R., Mental and Physical Measurements of Working Children, PSYCHOL. MONOG., 1914, No. 97, p. 67.

calculated on the basis of the number written by the superior or practiced hand. Starch's directions for a two-minute test of speed of writing was followed for both hands. The test was given twice and an average taken. These records are also included in Table I. The percentages vary from 29.7 to 74.8, the latter made by Aw who is naturally left-handed, but writes with the right hand. Some striking discrepancies occur with the records on the steadiness test. The speed records of the inferior hand are, however, subject to considerable correction when quality of writing is also taken into consideration.

One of us who had had considerable experience in rating handwriting attempted to arrange the left-hand specimens in an order of merit, with concealment of the names of the penmen. This arrangement was not felt to be highly satisfactory, particularly in the middle range. It was, however, quite possible to get a *group* arrangement that was fairly satisfactory. The specimens fell into seven groups. Specimens falling in groups I. and II. represent excellent left-hand writing. Table I. gives the quality of left-hand writing-quality for each subject. In a number of individual instances there is confirmation of the steadiness test. Lo, Ro, Bo, Ab, Da appear to be strongly right-handed. As is strongly left-handed; Wa and Dv are less so. Ambidextrality is indicated for DJ, Gr, OW, and DD on the basis of either speed or quality of left-hand writing. It is interesting to note that OW is father to AW; and DJ and DD are sisters.

Our method was subject to criticism in that our instructions did not discriminate carefully between speed and quality. In a repetition of this test with another group, we carefully discriminated in our instructions between the two tests, emphasizing speed in the first and quality in the second. The results were more satisfactory. We may remark in passing that we have come to expect ambidextral tendencies from penmen who write a back-slant or from those who show a tendency to shift from a running writing movement to a slow print-like type of script.

5. *Mirror-reading and Unidextrality.*—It remains to con-

sider the results of the three tests. Eighteen of our subjects were consistently right or left-handed; although with strong indications of a very great difference with respect to the extent of difference between the two hands. Seven subjects gave either inconsistent results or showed very slight difference between the two hands, namely, OW, DD, DJ, Wa, Sa, McG, Hi.

TABLE I.

Reagent	Right- or Left-Handed	Relative Unsteadiness of Inferior Hand	Relative Strength of Inferior Hand	Relative Speed Writing, Unpracticed Hand	Quality of Writing, Unpracticed Hand	Average Rank in Mirror Read	Difference in Length of Arms in 1/16 In.
Lo.....	R.	350	85.5	29.7	V	2	
Wa.....	L.	329.5	105.6	45.5	III	9.5	+ 1R
As.....	L.	321.1	92.2	39.1	II	13	+ 8L
Bo.....	R.	293.4	98.6	35.0	IV	16	+ 4R
Cl.....	R.	251.5	97.3	50.8	VI	5	+ 5R
Ro.....	R.	241.8	88	39.9	V	1	
Co.....	R.	220.1	90.6	46.3	VI	4	
McG....	R.	217.5	124.7	35.8	V	17	
Ab.....	R.	215.6	96.6	33.2	III	14.5	+ 4R
Dr.....	L.	214.1	94.7	47.3	II	11.5	
Wl.....	R.	210.2	59.1	41.6	III	19	
Sp.....	R.	206.1	99.1	46.4	VII	14.5	+ 1R
Da.....	R.	202.7	98.6	33.3	VI	7	+ 1R
Di.....	R.	197.3	98.5	28.9	V	9.5	
AW.....	L.	175	89.6	74.8	II	22	+ 1.5L
OW.....	R.	155	99.4	51.7	VI	25	+ 1R
Gr.....	R.	152.2	89.4	53	III	8	
DD.....	R.	146.3	107.1	36.3	I	21	+ 1R
Sa.....	R.	139.9	105.8	33.4	V	3	
Pa.....	R.	125.1	92.9	33.9	IV	6	+ 5R
Pe.....	R.	112.1	80.3	37.9	III	23	+ 2R
Ba.....	R.	106.3	85.1	34.4	IV	19	+ 2R
McK....	R.	103	97.3	36	IV	11.5	
DJ.....	R.	101	93.7	57.4	II	19	+ 3R
Hi.....	R.	55	95.1	37.5	V	24	0

Table I. gives the average rank of each subject for the mirror-reading tests. Lo and Ro, two of the most consistently right-handed of the group are found to excel in this test; OW, DD, DJ, McG, and Hi, five of the seven of the inconsistent group, make especially poor records.

Calculating by the formula,

$$\rho = 1 - \frac{6\Sigma D^2}{\eta(n^2 - 1)}$$

our coefficients of correlation for skill in mirror-reading and these various indices of dextrality, we get the following:

Mirror-Reading and Steadiness ..... .47 (P. E. .11)
Mirror-Reading and Strength of grip ..... .03 (P. E. .14)
Mirror-Reading and speed and quality of writing unpracticed hand .. .293(P. E. .12)
Pooling the tests for handedness by taking the algebraic sum of the deviations from the median gives a correlation with mirror-reading of ..... .51

The experiments with this series of twenty-five subjects give, then, some indication that skill in mirror-reading is correlated with degree of unidextrality. The evidence is cumulative and positive.

We were anxious, however, to secure better control of conditions, first in the way of a more thoroughly standardized test for mirror-reading and, secondly, in a more adequate and, if possible, simpler determination of degree of handedness. Accordingly, in 1916–1917, we approached the problem from a somewhat different standpoint.

Part II.

1. *Standardization of Mirror-Reading Test.*—In attempting to standardize the mirror-reading experiment we worked our two series of nonsense syllables, twenty-one syllables in each series (sixty-three letters), so constructed that every consonant appeared twice, once at the beginning and once at the end of each syllable and in combination with different vowels. These series of nonsense syllables, written in formal script, were etched and printed on standardized record-blanks with instructions for the test on the back of the blank. Test I. was given with a time limit of two minutes; Test II. with a time limit of one minute thirty seconds. Test II. was given immediately after Test I. Scores were calculated in terms of the number of letters correctly written in the time limit. A group of fifteen to twenty reagents could be tested at one session.

The reliability coefficient for Test I. and II. calculated from the scores of fifty-two freshmen and sophomores is .77 (P. E. .057). The correlation between these two tests is less high than expected. The instructions for the test are not wholly satisfactory. Some error arose, furthermore, from the group-testing and some from the fact that our

time limits were too long and that, consequently, in the case of subjects who finished before the signal it was necessary to record the time and calculate their score. These calculated scores might be a source of error in making an order of merit but are probably sufficiently accurate for the way in which we used them.

We tested, altogether, 180 subjects, 92 women and 88 men. Table II. shows the distribution of scores (see below).

TABLE II.

SCORE: NUMBER OF LETTERS CORRECTLY READ. MIRROR-READING I AND II

	1–9	10–19	20–29	30–39	40–49	50–59	60–69	70 and Over
I								
Men	1	6	15	17	24	14	2	9
Women	5	5	18	20	14	15	5	10
Total	6	11	33	37	38	29	7	19
II								
Men	2	6	9	5	22	22	11	11
Women	2	3	6	14	26	19	12	10
Total	4	9	15	19	48	41	23	21
Left-handed								
I	0	1	3	8	4	3	0	5
II	0	1	4	2	6	7	1	3
Right-handed								
I	6	10	30	29	34	26	7	14
II	4	8	11	17	42	34	22	18

2. *Mirror-Reading and Left-handedness.*—Having standardized the test in mirror-reading, we desired next some unequivocal method of determining degree of unidextrality. In this connection it occurred to us that studies on the inheritance of left-handedness give some indication that left-handedness is a Mendelian recessive and suggest the possibility not only that all left-handed subjects are pure in type but also that there is an "imperfection of dominance" shown in the existence of ambidexters.[1] If these suppositions were true, then on our hypothesis of a relation between skill in mirror-reading and unidextrality, a left-handed group tested for mirror-reading should show a higher central tendency

[1] Jordan, H. E., The Inheritance of Left-handedness, *Am. Breeders Mag.*, 1911, 2, 19–29 and 113–124.

than a right-handed group. Our group of 180 reagents contained twenty-four persons who claimed to be left-handed, although twelve of them write with the right hand. Table II. makes possible a comparison of the records of the left-handed and the right-handed group. There is certainly very little difference in distribution of scores. The slight difference that exists is in favor of the right-handed.

The results were disappointing but not surprising. From our previous work we had been inclined to believe that left-handed individuals show considerable variation in degree of unidextrality. For one thing, as we have mentioned before, they resist to an unequal degree the attempt to make them write with the conventional hand.

3. *Brachiometer Tests.*—While we were in the midst of this work, our attention was called to Professor W. F. Jones's method of determining congenital right and left-handedness by measurement of (1) length of the ulna, plus hand to the middle knuckle of little finger; (2) length of the humerus, (3) circumference of palm and (4) circumference of wrist. We immediately secured a brachiometer from Professor Jones together with his typed directions for taking the measures and a brief summary of his conclusions. Our information with regard to Professor Jones's work is limited to the typed pages sent out with the brachiometer and the report made by him at the 1916 joint meeting of the Psychological Association and Section L of the A. A. A. S.[1]

Very briefly, it is Professor Jones's contention that born-handedness is determinable by the measures cited above; adopted-handedness is shown by the measurement (1) of relaxed forearm circumference, (2) by contracted forearm, (3) relaxed arm (biceps) circumference, and (4) contracted arm (biceps) circumference.

The bulletin giving Professor Jones's full report upon the 10,000 measurements of paired bones made by him has not yet, we understand, been published, so that we are in the dark as to the detailed outcome of his investigation. In

[1] Jones, W. F., The Problem of Handedness in Education, PSYCHOL. BULL. 1917, 14, p. 64.

some quarters the question will be raised whether the difference in measurement of paired bones is the cause, the effect, or a concomitant of handedness. That difference in bone length is not the result of handedness would seem to be indicated by measurements on still-born and newly-born children and by the existence of the 'shiftover' or 'transfer,' the individual who uses (inexpertly according to Jones) the shorter arm. Jones's interest in the problem is largely the pedagogical one of the discovery and treatment of the 'shiftover.'

Our interest was somewhat different; we desired a way of classifying our subjects with reference to degree of handedness. So far as we know, Jones has not been particularly interested in this aspect of the problem, namely, the amount of difference in the two arm lengths and the possible correlation of greater difference with more pronounced unidextrality —ambidextrality resulting as an accompaniment of a vanishing difference in length. Questioned on the occasion of his report before the Psychological Association as to his understanding of ambidextrality, Professor Jones answered that the 'Ambidexter' is the 'shiftover.'

Whatever the interpretation of the figures, it was highly interesting from our point of view to measure the arm lengths of our reagents. Since we found it impossible to measure all our subjects, we selected for measurement those who belonged in the best and worst quartile in the mirror-reading test. Eight of our reagents were no longer accessible. This cut our total group down to 172, with forty-three subjects in each of the two groups to be measured. Furthermore, we measured all but five of our left-handed subjects and all who were accessible of the group of twenty-five with whom we had worked in the preceding year. All together we measured some one hundred persons. Limitation of time—both on our part and that of our subjects—prevented us from taking the measurements both for born and for adopted handedness; we limited ourselves to the four bone-measurements given above, the 'ulna plus,' humerus, circumference of palm, and of wrist.

It is perhaps necessary to observe that there is a considerable margin of error in taking these measurements with the instrument used. When the difference in length between the right and left forearm is, for example, only one sixteenth of an inch a high degree of accuracy may be questioned. In our tabulation these inaccuracies would in a measure cancel one another since they would occur equally in the measurements of both groups and lie in both directions. In order to insure as great consistency as possible, all measurements, with a few exceptions, were made by the same experimenter (Mr. Payson) who, following Jones' directions carefully, developed considerable expertness in taking the measurements, each of which he verified. Another advantage in having all measures taken by one person was that a bias as to the outcome of the individual measurements was avoided since the grouping into quartiles on the basis of the mirror-reading records was made by the other experimenter. We feel, on the whole, that the measurements of the 'ulna plus' are more accurate than those of the humerus; the determination of the wrist circumference more accurate than measurement of palm circumference. It was also easier to get a satisfactory record with the men, particularly in measuring the length of humerus, than with the women.

For a majority of the persons tested the four measurements were consistently in favor of the right (or left) arm and hand; in some 39 cases inconsistencies, often very slight, were recorded. For example, the right 'ulna plus' might be longer than the left; but the left humerus longer than the right. When such inconsistencies occurred the measurements were taken a number of times; sometimes the experimenter satisfied himself as to inconsistency; at other times the difference seemed to lie within the range of error. We shall await with great interest Jones's figures on this point.

With these limitations in mind we may turn to our tabular summaries (Tables III. and IV.). We give in detail the distribution for difference in length, 'ulna plus,' for the two arms and the median and central fifty per cent. for the difference in measurement for each of the four bone-measures taken.

TABLE III.

DIFFERENCE IN MEASUREMENT MAJOR AND MINOR ARM—BEST AND WORST QUARTILE (43 SUBJECTS EACH) MIRROR-READING I

Unit—one sixteenth of an inch for 'ulna plus'; one twentieth for wrist.

'Ulna plus' difference	0	1	2	3	4	5	6	Percentile difference
Best quartile (43)	4	6	9	9	8	3	4	.012
Worst quartile (43)	7	7	12	9	4	2	2	.009
Total (86)	11	13	21	18	12	5	6	
Wrist difference	0	1	2	3	4	5	6	
Best quartile (43)	15	8	12	3	2	1	2	
Worst quartile (43)	17	9	14	1	0	0	0	
Total (86)	32	17	26	4	2	1	2	

TABLE IV.

DIFFERENCE IN MEASUREMENTS MAJOR AND MINOR ARM—BEST AND WORST QUARTILE (43 SUBJECTS EACH). MIRROR-READING I

Units—one sixteenth and one twentieth of an inch.

	Best Quartile		Worst Quartile	
	Median	Range Central 50 Per Cent.	Median	Range Central 50 Per Cent.
'Ulna plus'	3	2–4	2	1–3
Humerus	2	1–4	1	0–4
Combined 'ulna plus' and humerus	4	3–6	3	1–5
Wrist	1	0–2	1	0–2
Palm	3	1–4	2	0–2
Cumulative measure	8	3–12	5	3–8

The difference between the two arms for the 'ulna plus' and the humerus is given in terms of one sixteenth of an inch, the unit of the brachiometer; and for wrist and palm in terms of one twentieth of an inch, the unit of the anthropometric tape. The combined measure for the 'ulna plus' and the humerus is obtained from the algebraic sum of the two measurements. The cumulative measure is simply an arbitrary index of the difference between the two arms and hands of the subjects; we took an algebraic sum of the four measures for each individual, ignoring the fact that the unit was one sixteenth of an inch in the ulna and humerus measurements and one twentieth for wrist and palm, and ignoring also the fact that the difference occurs on a much larger whole in the first two than in the second two measures.

We may meet at this point a criticism that would naturally

be made, namely, that difference in the right and left measurements should be given not in absolute terms but in terms of the relative proportion to the whole length measured. Actually, however, a variation of four-sixteenths of an inch on 17 inches and on 13 inches (extreme measurements for the 'ulna plus') shows less difference than a variation from one to four-sixteenths on either of the two given measurements. Some inaccuracy is, however, incident to giving the figures in an absolute rather than a relative way. We calculated, therefore, for the 'ulna plus' the value in percentage terms, that is, in terms of the proportional difference for each member of our group of eighty-six and obtained the average for each of our two groups; for the best quartile the difference is .012 in terms of the total length of the 'ulna plus'; for the worst quartile, .009, a slight reduction on the median as given in terms of the absolute difference. For all practical purposes the absolute difference is nearly as accurate.

A study of Tables III. and IV. indicates that with the exception of the wrist measurements, the best quartile shows a slight but consistently greater difference between the right and left hand and arm than does the worst quartile. The complete distribution of the wrist measures also shows a slight excess in favor of the best quartile. The outcome is interesting and suggestive if nothing more. Taken in conjunction with other evidence, it certainly points to a significant problem and perhaps indicates a definite conclusion.

On the other hand, a number of striking exceptions to our general conclusions occurred. Close study of individual cases would probably repay the effort required.

It is at least interesting to scrutinize certain groups in the light of the measurements. In the last column of Table I., we give the combined 'ulna plus' and humerus measurements for the fifteen subjects of the previous investigation who were available for the brachiometer measurements. So far as these measurements go they are largely confirmatory of the conclusions drawn from other tests. Reagent As, for example, whom we had on other grounds selected as the most left-handed of this group of subjects shows a more con-

siderable difference in favor of the left arm than any other left-handed subject, with one exception, that we measured. Aw and Ow (son and father) show a very slight difference in the length of the two arms, for Hi the difference is imperceptible. We have already cited the coefficient of correlation for mirror-reading and degree of unidextrality as determined by the steadiness test.

We may remark here that this group should have been more rigidly selected on the age basis, a failure which probably caused a margin of error in three or four records, particularly Bo's and Pe's, who are consistently right-handed but older than the majority of the group. The other two older reagents (OW and JD) show strong ambidextral tendencies. In Part II., also, we have ignored the age factor. But in passing, we may suggest that the relation of degree of unidextrality to age remains an outstanding problem. Various records have been given by different investigators that point to an increase of handedness at puberty. It is possible that a more extensive investigation would repay the work necessary and possibly show some interesting epochs.

Another group of subjects that we need to examine consists of the twenty-four left-handed reagents. We were able to get the measurements of nineteen of these. One of the group proved to be what Jones calls a transfer or shiftover since her right-side measurements were consistently larger than the left-side ones. She herself is very dubious as to our interpretation and says she was not shifted by an accident, so far, of course, as her knowledge goes, and has no inclination to be right-handed. Her mother is left-handed and it is possible that the left-handedness of the daughter may be explained by imitation.

The other eighteen left-handed persons show considerable variation in the amount of difference between the two arms. It varies from 0 to 9, for the combined arm length. The extreme measurement is less than our extreme in the right-handed group but this would be anticipated from its smaller number. Eight of the eighteen write with the minor hand, in this case the right hand. With two exceptions this group

of eight shows only a slight difference in the measurements of the two arms, and, in general, a less difference than that found for most of the subjects of the other group. The exceptions in the first group are its two oldest members. Parental and pedagogical pressure in forcing a shift from the left to the right hand in writing is becoming milder. Undoubtedly under present school-room conditions these two individuals would be writing with the left hand.

About four out of a hundred individuals are born left-handed according to Jones but three of them are using the minor hand. Among the one hundred supposedly-right-handed persons whom we measured we found three who gave consistently for the four measures an excess in favor of the left side. One of the subjects reported himself as ambidextrous, and in fact the difference in favor of the left is very slight. He is very inexpert at mirror-reading. The other two belong in our best quartile for mirror-reading. One of them admitted that he was left-handed. He calls himself right-handed because of writing with the right hand; he does, however, a number of things with the left hand. No particular pressure was exerted to make him write with the minor hand. As a matter of fact the difference between the two hands is not extreme. The third case shows an extreme difference between the two arms and hands. Questioned in detail it is obvious that this subject does many things with his left hand. He supposed this was due to accident in forming certain habits. After his attention was called to the matter he reported a considerable number of situations in which he used the left hand in preference to the right.

Besides these three reagents seven others gave an excess of difference in favor of the left arm if we take the combined measures of the 'ulna plus' and humerus; their records, that is, are inconsistent. The difference, however, is very slight except in two instances. Five of the seven are very inexpert mirror-readers. This group would be a most interesting one to study.

4. *Puzzle-tests.*—Throughout our work on mirror-reading the impression has persisted that facility in such reading is

merely a special instance of skill in handling spatial relationships. The puzzle-tests to be described in this section of the paper were designed for further investigation of the nature of this ability. It seemed probable that one of two things was true. Either unidextrality was connected with a general ability to appreciate all sorts of form relationship or it was confined to an understanding of symmetry. Two kinds of puzzles were therefore designed, one of them laying emphasis on the symmetrical relationship and the other requiring, instead, an aptitude for the manipulation of form. Three puzzles of each kind were used. Our subjects were twenty-five freshmen.

Series I. Symmetrical Puzzles.—The puzzles in this group vary somewhat in appearance and character but exhibit in the main the same fundamental principles. To each of these puzzles there were two parts; a diagram and a dissected symmetrical counterpart of the diagram. The pieces of the diagram were numbered and the pieces of the symmetrical part removed and placed on the table beside the subject. The subject was required to place the pieces *in the order in which they were numbered on the diagram.* To do this it was necessary for the subject to refer constantly to the diagram and make the translation into its symmetrical form. It was thought that this ability was very similar to the ability to translate symmetrical writing, that is, mirror script. The following directions were given before the puzzles were shown to the subject.

"Do you understand the difference between symmetrical and similar figures? Symmetrical figures have essentially the same form but the parts are reversed in such a way that they may not be fitted together. Your two hands, for example, are symmetrical.

"I am going to give you some puzzles to fit together that involve the principle of symmetrical figures. You will have a diagram to guide you which is symmetrical to the figure you will fit together and so you will have to keep the difference between symmetrical and similar figures always in mind or you will become confused.

"The pieces are numbered on the diagram in the order in which they are to be placed. Do not turn them over."

When it was certain that the subject understood the directions, the puzzles were produced. The parts were placed on the table at one side of the puzzle in a given order, so chosen as to give the subject no hint as to the solution. The time required to place each piece was taken and an order of merit obtained based on the time required to do the entire puzzle.

As previously mentioned the symmetrical puzzles differ somewhat from one another and a brief description of each is given below.

Symmetrical Puzzle No. 1.—This puzzle is rectangular in form and is composed of six differently shaped figures bounded by irregular curves and straight lines. It is, perhaps, the easiest to handle of the three. The pieces are numbered in an irregular order.

Symmetrical Puzzle No. 2.—This and the third puzzle are rather similar in design. Both are circular in outline and both composed of parts radiating from a central point. Puzzle No. 2 is formed of rather similar parts numbered consecutively. In this puzzle it is particularly difficult to recognize the separate parts from inspecting the symmetrical diagram. Anyone not possessing a keen perception of form might be expected to have difficulty with this puzzle.

Symmetrical Puzzle No. 3.—This puzzle differs principally from the preceding one in having parts that are different enough from one another to be more easily recognized by aid of the symmetrical diagram. It is probably less a test of form perception. The pieces are numbered in an irregular order.

The method pursued in these tests—which involved the placing of the dissected pieces in the order indicated by the number on the diagram—eliminated one element of variability in individual manipulation of puzzle pieces and made it possible to record the time needed for placing each insert, a necessary preliminary for standardization of such a test. It did not, however, overcome blocking and possibly increased

it since free trial and error was in a certain measure ruled out. A possible improvement in method might be made by limiting the time for placing of each insert. Some method of quantitative standardization seems necessary if we are to utilize puzzle tests in any other than a purely individual way. Many authorities assert, however, that the method of handling a puzzle is throughout much more significant than is the quantitative factor.

Series II. Angular Puzzles.—The first puzzle of this series is that known as the Kempf-Knox Diagonal Test. The two others are somewhat similar to it in design but more difficult. They are composed each of six triangular pieces and are so constructed that three of the pieces, if properly placed, form one half of the puzzle exactly symmetrical to the other half. It was thought that a good subject would realize that the pieces were in pairs, solve half of the puzzle and then easily construct its symmetrical half. As a matter of fact very few, if any, of the subjects made use of this line of reasoning; consequently, the puzzles were as difficult as if they had been composed of six dissimilar pieces. The instructions given were extremely simple. The pieces were spread out upon the table and the subject told to fit them into the frame. The only record kept was the total time required to complete each puzzle.

In criticism of these angular puzzle-tests it may be said that they proved to be too difficult. A series as simple as the Kempf-Knox Diagonal Test would be preferable. Blocking on a single piece may interfere with the record in such a way that a poor subject happening to find a good combination may make a much better record that a good subject not so fortunate in the placing of the first pieces. With all these imperfections in method, individual differences are plainly evident. Some subjects are very evidently good at manipulating the blocks so as to fill a given space while others are just as evidently confused.

To return to our original question. An order of merit based on the average time required for solving the symmetrical puzzles correlated with that based on two tests for

mirror reading .43 (P. E. .11). An order of merit based on the average of the Angular Puzzles No. 2 and 3 gave a correlation of .15 (P. E. .13). These figures offer some evidence of relationship between the mirror-reading test and the solving of symmetrical puzzles but are inconclusive so far as angular puzzles are concerned. Many expert mirror-readers are so obviously superior in their fitting together of puzzles that we feel that investigators of puzzles tests would do well to keep in mind unidextrality as a possible factor in the case.[1] No doubt, close study of our individual cases that represent discrepant results would yield valuable material.

5. *Mirror-Reading and General Intelligence.*—Mirror-reading is, we believe, a highly specialized capacity. It is a common occurrence to find surprising expertness or inexpertness in unexpected quarters. As a side-light on this aspect of the situation it is interesting to cite the correlational coefficient between an order of merit based on skill in mirror-reading and one based on the Intelligence Quotient obtained by use of the Stanford Adult Tests. I. Q's were at hand for forty of the subjects tested in mirror-reading; the correlational coefficient was .14 (P. E. 11). This coefficient does not, of course, prove that general intelligence is of no value in adaptation to the test. With a small group such a coefficient is indecisive but at least it weighs against the possibility of a very high correlation. The outcome is, moreover, in agreement with the failure to find any indication of a correlation between mirror-reading and the completion test, as reported in a former paper.[2]

6. *Mirror-Reading and Visual Capacity.*—Throughout our work on mirror-reading we have had a strong feeling that some connection exists between this and visual capacity (visual imaging, visual attention and the like). Certainly we are collecting an increasing number of clever abstract thinkers, often literary in type of mind, who are hopelessly

[1] The twelve superior subjects in both sets of puzzles gave an average of arm-measurement in excess of that of the twelve inferior subjects. But the subjects in each group are so few and the M. V. so large that the average has only suggestive value.

[2] Downey and Anderson, *loc. cit.*, p. 360.

adrift when confronted with mirror script, and of the slower visual type who handle it with perfect ease. The coincidences seem too frequent to be accidental. It is, of course, quite possible for one's observations to become biased; and in the absence of an objective test for visualization one cannot discuss the problem with any satisfaction. We attempted a study of two small groups of freshmen, six in each, chosen from the two extremes in mirror-reading and exhibiting in the one case considerable difference in the bone-measures and very slight difference in the other. Four of the latter, or poor, group are very deficient in visual control as evidenced by their scores and their introspective reports in the clock test (Stanford revision, XIV. 6) Code Test (Stanford revision, XVI. 6) and the Cube Test (Langfeld and Allport)[1] but two of this group seem to handle visual imagery with ease. Of the good group four give evidence of excellent control over their visual imagery, one is a fair visualizer, one is somewhat deficient. The relationship is one that deserves detailed investigation not only because of the light it may throw on the explanation of mirror-reading but also because of its possible complication with unidextrality.[2]

7. *Summary and Conclusions.*—As an outcome of our work we are convinced that the proposition that asserts a relationship existing between mirror-reading capacity and unidextrality deserves serious consideration. Furthermore, skill in mirror-reading appears to be indicative of a specialized capacity which is of value in some forms of spatial orientation. There is also a possibility that mirror-reading and visual capacity are related abilities. The real point of interest in all this is not, of course, the citing of factors concerned in mirror-reading as mirror-reading. *This is incidental to a bigger question whether or not unidextrality, or rather degree of unidextrality, is bound up with certain mental attributes that, taken together, constitute a somewhat definite mental type.* In

[1] 'An Elementary Laboratory Course in Psychology,' p. 72.

[2] We have at hand some evidence of correlation between speed of mirror-reading and speed in reversed spelling of words (a visual test). The latter test *if standardized* would be of considerable value.

the absence of evidence suggestions must be recognized as purely speculative.

The real antithesis appears to lie between ambidextrality and unidextrality and not between right- and left-handedness. Possibly there are two factors involved, one governing the degree of dominance and the other placing it. Close study of family groups should be instituted. Quite likely we should find the very right-handed and the very left-handed persons in the same family.[1] To make our records of value we must have some way of checking up reports. If large groups are to be studied this method must be a comparatively simple one. It is for this reason that Jones's measurements seem to us so promising, and that quite apart from the questions concerning the causal or indirect relationship existing between handedness and difference in length of paired bones. At present we are pretty much in the dark as to the anatomical basis of unidextrality. In particular, we should like to know more of the cerebral conditions found to exist in the case of individuals of pronouncedly unidextral or ambidextral tendencies. Here we must await neurological investigation.

Meanwhile, there are a number of investigations that need reviewing with the possibility in mind that approximate ambidextrality and extreme unidextrality may be correlated with somewhat definite mental differences. One of these reviews should consider the evidence that seeks to relate speech defects to shifting from the natural hand in writing. Certainly we have need for more precise data both with reference to the kind and extent of speech difficulty and the degree to which a dominant-handedness actually existed. Again, evidence that mental defectives are predominatingly left-handed or ambidextrous should be reviewed in connection with more adequate measurements of what constitutes dextrality, sinstrality or ambidextrality. Doll,[2] in harmony with other investigators, observes that "sinistrality in grip does not necessarily imply sinistrality in other manual activities"

[1] Should extreme unidextrality be found correlated with a determined mental type, we should have in an index of unidextrality a measure of more than handedness, and the way would be open for enlightening studies on human inheritance.

[2] 'Anthropometry as an Aid to Mental Diagnosis,' p. 65.

and yet often the judgment as to the dominant hand is based upon the relatively stronger grip. In particular, it might be worth determining whether mental defectives show striking variations on the manual side and on the verbal that might possibly serve to place them in groups. If so, the possible unidextrality of the first (or objective-minded) and the approximate ambidextrality of the second (or subjective-minded) group would be an interesting question for investigation.

THE DIFFERENTIAL SPATIAL LIMEN FOR FINGER SPAN

BY HERBERT SIDNEY LANGFELD

Harvard University

The investigation of the limen for finger span was suggested by the practical demand for accuracy in estimating the number of objects or size of objects grasped between the thumb and index finger. The problem of psychophysical measurements in general, however, although one of the oldest, has rather increased than diminished in theoretical interest and importance, as the many functions involved, and at first so little suspected, have come to light, thus casting suspicion upon the accuracy of threshold measurements. In representing these results, therefore, the theoretical question of method and the desire to present further data for discussion are uppermost.

Beside the establishment of the limen, the experiment was arranged to ascertain individual differences, differences according to the method of presentation of the stimuli, the relation of the subjective feeling of certainty to the objective accuracy, and the value of the doubtful or guess judgments.

There were six subjects—five men and one woman. Three of them, *A*, *B*, and *F*, were trained experimentalists, of which *B* was the best trained. The other subjects had little knowledge in methods. Subjects *A* and *B* were slow, deliberate, and careful. The others were quicker and of a more nervous temperament. Subject *E* was particularly high strung and nervous to the extent of bodily twitchings. This description was corroborated by another experimenter who had all but one of these subjects.

The instruments used were six calipers, two of which were Starrett's No. 25 M as shown in the cut, page 419. The other four were modelled from them. Small, accurately machined right-angle plates with perpendicular surfaces,

16 mm. by 25.5 mm., were securely screwed on the ends of the two points (*A*). At the end of the adjustable arm rest were upright blocks which held the clips (*B*) into which the calipers could be easily slipped. The arm rest was adjusted for each subject so that when the hand rested comfortably on the board with fingers extended over the edge, the caliper plates could be grasped in their middle without moving the hand.

The procedure adopted was the method of constant stimuli. Not only did this method seem less open to theoretical criticism, but it was more practical for this experiment than the method of limits, since it would have been difficult to obtain an instrument that could be quickly adjusted with the necessary degree of accuracy. The standard distance was 5 cm., a distance arbitrarily selected as not being too great to be comfortably spanned by all the subjects and yet sufficiently large that relatively large differences could be used. The comparison stimulus was set at a distance of 1 mm. Five comparison stimuli were used, the two extremes being eliminated.[1] There were, then, two standard stimuli of 5 cm. each and four comparison stimuli of 5.2 cm., 5.1 cm., 4.9 cm., and 4.8 cm. respectively.

A long series of about seven hundred judgments for each subject was run through during the first half year and the effect of practice was thus in great part eliminated.[2] The results have not been included in this paper, and the method employed, which differed from that finally adopted, need not be described here.

The main experiment followed three weeks after this practice series (an unfortunate break due to the mid-year period). Three separate limens were obtained: one for the span of the index finger and thumb of the right hand, the standard and comparison stimuli being presented successively

[1] See Fernberger, S. W., On the Elimination of the Two Extreme Intensities of the Comparison Stimuli in the Method of Constant Stimuli, PSYCHOL. REV., 1914, 21, 335–355.

[2] Fernberger, S. W., The Effects of Practice in its Initial Stages in Lifted Weight Experiments and its Bearing Upon Anthropometric Measurements, *Amer. J. of Psychol.*, 1916, 27, 261–272.

to the same hand; another for the two hands, the stimuli being presented simultaneously to the two hands, the standard to one and the comparison to the other; a third, also for the two hands, but with the stimuli presented successively, first the standard to the one hand followed by the comparison to the other, and vice versa.

The subjects worked in pairs, alternating after each set of thirty judgments in order to prevent fatigue, and only ninety judgments were made in the hour. A week intervened between each group of ninety judgments. Each set of thirty judgments consisted of ten judgments for each of the three limens. The subject made ten judgments for the limen of the right hand, then ten when the stimuli were presented simultaneously to the two hands, and finally ten when the stimuli were presented successively to the two hands. On half the days this order was reversed, the stimuli being presented first to the two hands successively, and last to the right hand alone. The subject then read while the other subject went through the thirty judgments. The series of ninety judgments for the two subjects lasted about an hour. The entire experiment extended through the four months of the second semester.

The order of the presentation of the comparison stimuli was 5.2 cm., 4.9 cm., 5 cm., 5.1 cm., and 4.8 cm., with the comparison stimuli presented first in the successive presentation and to the right hand when two hands were used. This same order for the comparison stimuli was then repeated with the standard first and to the left hand when two hands were used, the time and space error being thus objectively eliminated. The judgment was given in terms of the second stimulus when successive and in terms of the right hand stimulus when simultaneous. The stimuli were applied at intervals of a second.

The subject sat at a convenient height in front of the table with his hand lying comfortably upon the rest, which was properly adjusted so that the fingers could grasp the plates at their centers. All the calipers were so set that the two plates were equidistant from the middle point of the

instrument. In thus keeping the center of the span always in the same place, there was no possibility of detecting a difference through a change in the symmetrical relation of the two plates to the hand. The subject closed the second, third, and fourth fingers and extended the thumb and index fingers, which were held a convenient distance apart. He closed his eyes, the experimenter placed the calipers and clips, and said 'Now.' The subject then closed his fingers on the two plates and gave his judgment. He was never at any time given a clue as to the accuracy of the results.

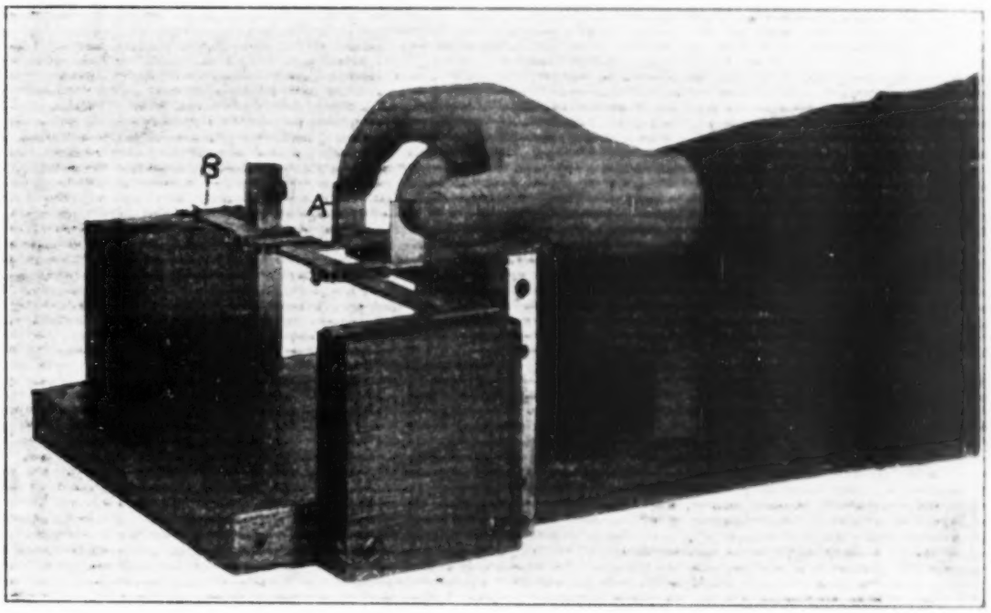

FIG. 1. Arm Rest with Caliper and Hand in Position.

The subject was instructed to clasp the plate with sufficient pressure to be able to feel the spatial extent separating them, the sensations involved being principally those of joint and tendon. He was particularly cautioned not to use the distance moved through by the fingers in clasping as a criterion. Special attention was paid to practice in abstracting from this movement perception during the five months of preliminary training, and from the introspection of all the subjects, it is fair to assume that there was the proper direction of attention and that the judgment was based on the static spatial perception.[1] The subjects were told to judge

[1] Inasmuch as the fingers did not always assume the same position previous to the grasping, this criterion of movement would have been a very uncertain one.

whether the second space was larger, smaller, or equal to the first, or in case of simultaneous presentation, whether the right was larger, smaller, or equal to the left. He was further told that there were objectively equal stimuli. He was directed to say equal only when there was an actual perception of equality, and not when there was a mere perception of no difference. He was directed to grade his judgments *a*, *b*, and *c*, according to his degree of confidence in his answer, *a* being the highest degree. If he could not make a larger, smaller, or equal judgment even with *c* confidence, he was to say 'equal *d*,' and was then to guess either larger or smaller. Under this category would fall all doubtful judgments, no difference judgments, and or-judgments.[1] Both judgments were noted for separate treatment. The response under these instructions, for example, took the form of 'larger *c*,' 'smaller *a*,' 'larger *b*,' etc. None of the subjects had the least difficulty in grading their confidence, the confidence response coming immediately and without hesitation after the first judgment. Even during the first series, no one complained of it interfering with his judgment, and as it did not take an appreciably longer time, it was decided to continue it throughout the series even though correlations of degree of confidence with accuracy of results and an analysis of the guess judgment had been made by previous investigators.

The method of treatment of the results was according to the $\varphi(\gamma)$ hypothesis.[2] The upper and lower limens, S_2 and S_1, the interval of uncertainty, $S_2 - S_1$, the Measurement of Sensitivity or Threshold of Volkmann, $(s_2 - s_1)/2$, subjective equality, $(c_2 + c_1)/(h_2 + h_1)$, and coefficients of precision, h_1 and h_2, were calculated. The results are recorded in Table I.

Directing our attention first to the M. S.[3] for the one hand,

[1] S. S. George believes that doubtful and or-judgments probably belong to the same category. For an excellent evaluation of these various judgments and an explanation of their non-serial nature, see his article entitled Attitude in Relation to the Psychophysical Judgment, *Amer. J. of Psychol.*, 1917, **28**, 1–37.

[2] See F. M. Urban, The Method of Constant Stimuli and Its Generalization, Psychol. Rev., 1910, **17**, 229–259. Also, S. W. Fernberger, On the Relation of the Methods of Just Perceptible Differences and Constant Stimuli, Psychol. Monog., 1913, No. 61, especially pp. 29 and 38. I am greatly indebted to Professor Fernberger for his valuable advice in regard to the mathematical treatment.

[3] 'Measurement of Sensitivity.'

it will be noticed that there is a considerable difference in the individual M. S. as well as in the coefficients of precision. The two subjects who have the highest coefficients of precision, subjects *A* and *B*, are two of the three best trained

TABLE I

Method	Subjects	h_1	h_2	$S_1=\left(\frac{C_1}{h_1}\right)$	$S_2=\left(\frac{C_2}{h_2}\right)$	S_2-S_1	$\frac{S_2-S_1}{2}$	$\frac{C_2+C_1}{h_2+h_1}$
Right hand	*A*.....	6.1604	5.6587	4.976	5.066	0.09	0.045	5.019
	B.....	5.520	5.042	4.94	5.03	0.09	0.045	4.987
		(4.9632)	(4.849)	(4.917)	(5.049)	(0.132)	(0.066)	(4.982)
	C.....	2.5922	3.7562	4.977	5.075	0.098	0.049	5.035
	D.....	3.5630	3.3988	5.009	5.074	0.065	0.032	5.044
	E.....	3.1709	2.9872	4.978	5.042	0.064	0.032	5.009
	F.....	(4.2877)	(4.5568)	(4.98)	(5.056)	(0.076)	(0.038)	(5.019)
Simultaneous	*A*.....	3.5461	3.978	4.982	5.081	0.099	0.049	5.035
	B.....	3.1252	2.809	4.822	5.097	0.275	0.137	4.952
		(2.9785)	(2.4428)	(4.772)	(5.135)	(0.363)	(0.181)	(4.935)
	C.....	1.3948	2.5185	4.966	5.002	0.036	0.018	4.989
	D.....	2.0402	2.5045	4.929	5.15	0.221	0.11	5.051
	F.....	3.9459	4.2674	5.025	5.035	0.01	0.005	5.03
	F.....	(1.6789)	(2.868)	(4.736)	(5.096)	(0.36)	(0.18)	(4.963)
Successive	*A*.....	4.0334	4.7442	4.97	5.087	0.117	0.058	5.033
	B.....	3.1591	2.412	4.894	5.102	0.208	0.104	4.984
		(2.4178)	(2.1034)	(4.821)	(5.155)	(0.334)	(0.167)	(4.977)
	C.....	2.5538	3.3941	4.933	5.047	0.113	0.056	5.002
	D.....	1.7494	2.6402	5.007	5.137	0.131	0.065	5.085
	E.....	2.6394	3.0165	5.02	5.068	0.048	0.024	5.046
	F.....	(2.0141)	(2.7383)	(4.951)	(5.058)	(0.107)	(0.053)	(5.012)

h_1 = Coefficient of precision "less judgments,"
h_2 = Coefficient of precision "greater judgments,"
S_1 = Lower limen,
S_2 = Upper limen,
$S_2 - S_1$ = Interval of uncertainty,
$\frac{S_2 - S_1}{2}$ = Measurement of sensitivity (Threshold of Volkmann),
$\frac{C_2 + C_1}{h_2 + h_1}$ = Point of subjective equality.

subjects, and the two who were the most careful and consistent in their attitude toward the experiment. Even in the other two methods, subject *A*'s coefficients are the highest, and *B*'s coefficients rank very high. Although some of the other subjects had lower M. S.'s, it seems preferable to consider the M. S.'s of these two subjects, which are identical, as the M. S. under the most constant conditions, rather than to average all the subjects. The data from the other subjects have been included, however, for comparison. The

M. S. is, then, 0.45 mm., which bears a ratio to the standard of about 1/100.[1]

The data of subject *F* have unfortunately little value and have therefore been bracketed. This subject and subject *B* had a great many doubtful judgments. For subject *B* there are available a sufficient number of judgments beyond the fifty required to permit the elimination of these doubtful judgments entirely and to substitute other judgments in their place. This was impossible in the case of subject *F*, so that the data include the doubtful judgments as equal. The data were calculated for subject *B*, both with and without the doubtful judgments included, the calculations for the former being included (bracketed) for comparison. It will be seen that with the doubtful judgments eliminated, the M. S. is lower and the precision higher. An explanation for this is given on page 427 where the doubtful judgments receive further treatment. Another factor that makes subject *F*'s results questionable is that during work in another university he had acquired a set not to give equal judgments and it was rather difficult for him to change this.[2] That may be a reason why he had so many doubtful judgments. His data are only used in discussing relative values. The other five subjects had very few doubtful judgments, so few that it was not necessary to eliminate them, as they made no appreciable difference in the results. The exact number of doubtful judgments for the subjects in all three methods, that is in seven hundred and fifty judgments, is as follows: subject *A*, 3; *B*, 100; *C*, 10; *D*, 14; *E*, 6; and *F*, 86.

The results of the subjective equality increase the confidence in the method. In all three methods of presentation and for all subjects, it approaches closely to objective equality. When we compare the results of the three methods of presentation, we find little consistency among subjects. From experience we would have prophesied higher M. S.'s and less

[1] Jastrow some years ago made some experiments upon finger span with wooden blocks of various sizes. He found that with a difference of 1/100 the subject made 3.5 errors in ten judgments. The Perception of Space by Disparate Senses, *Mind*, 1886, p. 552.

[2] See S. W. Fernberger, The Effect of the Attitude of the Subject upon the Measure of Sensitivity, *Amer. J. of Psychol.*, 1914, 25, 538–543.

precision for two hands. For precision this is the case, with the exception of subject *E* for simultaneous presentation. As regards the M. S., the subjects vary among themselves, even the two whom we have selected as most reliable. Subject *A*'s M. S. for the right hand is the lowest, but it is only slightly lower than his simultaneous, and not much lower than his successive M. S. Subject *B*'s M. S.'s, on the other hand, are much higher for simultaneous and successive presentations. Four subjects' M. S.'s are lower for the right hand, and two for simultaneous presentation. Some are lower for simultaneous than for successive presentation, and others the reverse. There are so many possible factors involved that it is impossible from the data at hand to offer satisfactory explanations. There is no doubt, especially with untrained subjects, that various criteria were used. Judging from subject *A*'s consistent results, it is probable that his attitude and his criteria of judgment remained fairly constant. But from what has recently been found in regard to the effect of attitude on limen,[1] it is useless to discuss individual differences in regard to sensitivity, unless one is sure that the attitude is constant or is cognizant of the nature of the change.

The ranking of the five subjects according to their M. S. for the three methods is:

Right hand	*E D (A B) C*
Simultaneous	*E C A D B*
Successive	*E C A D B*

There is perfect correlation between the two methods for the two hands. The change of rank appears in the change from one hand to two hands, the degree of correlation being only .18.

The exceedingly low, in fact very improbable M. S. of .01 mm. of subject *E*, makes one suspect that in his case fifty judgments for each pair were not sufficient. It seems too low even though his nervous jerky manner in holding the stimuli may have changed the simultaneous presentation into a quick and repeated alternation of stimuli.[2]

[1] See Fernberger, *op. cit.*, and George, *op. cit.*

[2] While on the subject of the number of judgments, it might be said that the decrease of all the coefficients of precision for the two-hand judgments is an indication that it would have been better to have made more judgments for these series.

Another point of interest is that the S_2 is further from the objective equality than the S_1, in practically all cases. This indicates the constant error or general tendency of judgment and the shifting here is in the direction of the larger judgments. If more subjects had been used, it is possible that there would have been those who preferred smaller judgments and those who showed no preference, that is, negative and indifferent types as well as positive types, such as Martin and Müller found with lifted weights.[1] There are certainly not sufficient cases to conclude that only the positive type exists in determining this limen, although this type seemed to predominate in most of the previous experiments. On the other hand, one of the explanations offered relative to the lifted weight experiment, namely, that the types depend upon the strength of the subject, is hardly applicable here, unless it happens that one develops a general set in the course of judgments depending primarily upon strength and that then this set influences all judgments of greater or less.

In Table II. are recorded in percentage all *a*, *b*, and *c* confidence according to right and wrong judgments. The figures are calculated from the data of all six subjects. When the $d = 0$, there is little difference in the per cent. of *a* judgments for the larger and smaller judgments. The figures are 21 per cent. compared to 17 per cent. But when $d = 1$ mm., many more correct answers have *a* confidence than incorrect, 20 per cent. and 25 per cent. correct to 8 per cent. and 9 per cent. incorrect. When $d = 2$ mm., this difference is still greater, 31 per cent. and 34 per cent. correct to 10 per cent. and 9 per cent. incorrect. The distribution of the *b* judgments is in the same direction, but the differences are not so marked.[2] It is evident that the mere judgments

[1] 'Zur Analyse d. Unterschiedsempfindlichkeit,' Leipzig, 1889. Fullerton and Cattell among others found that the judgments were more often right when the second weight was heavier. 'On the Perception of Small Differences,' University of Pennsylvania Publications, Philosophical Series, 2, 1892, p. 127. It seems impossible entirely to eliminate the time and space errors and obtain symmetrical curves. If other investigators had not found a similar lack of symmetry we should have thought we had perhaps too few judgments. Even so, fifty judgments seem the minimum and one hundred would be better.

[2] Pearse and Jastrow found a similar relation between confidence and correctness.

'greater' or 'less' do not give a complete picture of the effect of the difference. If we had used categories such as 'much greater' and 'much less,' probably the differences which here were expressed in a subjective scale would be expressed in the actual judgments. Suppose 'greater' was the correct judgment, then if conditions were the same as in this experiment, there would have been more 'much greater' judgments than 'much less' judgments, if these categories had been added.[1]

TABLE II

Com. Stim.	*a*	*b*	*c*	*a*	*b*	*c*	*a*	*b*	*c*
		<			=			>	
5	66	111	201	0	16	130	71	106	156
$d = 0$	17	30	53	0	10	90	21	32	47
	< (wrong)			=			> (correct)		
5.2	12	44	63	0	8	91	198	225	223
$d = 2$	10	37	53	0	8	91	31	35	34
	> (wrong)			=			< (correct)		
4.8	8	19	62	0	19	115	212	191	221
$d = 2$	9	21	70	0	14	86	34	31	35
	< (wrong)			=			> (correct)		
5.1	16	73	115	1	23	118	105	180	217
$d = 1$	8	36	56	0	17	83	25	25	50
	> (wrong)			=			< (correct)		
4.9	15	48	101	0	14	145	106	185	241
$d = 1$	9	30	61	0	9	91	20	35	45

a, *b*, *c* = degrees of confidence. The first horizontal column of figures for each comparative stimulus shows the number of judgments for each degree of confidence and for each class of judgments. The second horizontal column shows the per cent. of each of the three degrees of confidence for each class of judgments.

See Jastrow's A Critique of Psychophysical Methods, *Amer. J. of Psychol.*, 1887–1888, p. 305. Fullerton and Cattell also found that the degree of confidence varies as the percentage of right guesses. *Op. cit.*, p. 126. Harold Griffing, in his article on Sensations from Pressure and Impact, Psychol. Rev. Monog., 1895, p. 45, reports the same relation. As there were very few *a* judgments, this relation holds for the *b* judgments. G. M. Whipple, in An Analytic Study of the Memory Judgment and the Process of Judgment in the Discrimination of Clangs and Tones, *Amer. J. of Psychol.*, 1901, p. 446, also found much greater certainty for right answers.

[1] H. Ebbinghaus advised the use of the following categories: deutlich kleiner, ebenmerklich kleiner, gleich, ebenmerklich grösser, und deutlich grösser, 'Grundzüge der Psychologie,' Bd. 1, 1892, p. 74.

It is also interesting to note that the equality judgments were never given with confidence and the amount of *b* confidence is less than with the other judgments.

The subjects, ranked in order of confidence according to the number of *a* judgments, are: subject *E*, 311; subject *A*, 280; subject *D*, 129; subject *B*, 38; subject *C*, 37; subject *F*, 15 (see Table III.). It is seen that only subjects *E* and *A*

TABLE III

DISTRIBUTION OF THE JUDGMENTS OF EACH SUBJECT ACCORDING TO CONFIDENCE

Confidence	Subject *A*	Subject *B*	Subject *C*	Subject *D*	Subject *E*	Subject *F*
a.....	280	38	37	129	311	15
b.....	168	170	230	270	273	174
c.....	298	448	472	334	162	478

had a high degree of confidence. Subject *D* had a medium degree, and subjects *B*, *C*, and *F* very little. This assumes that they are using the same scale. Even though one subject's *a* means a higher degree of confidence than another's, however, it seems as if the number of *a*'s should give some indication of the subjective state. There is the usual lack of agreement between accuracy and confidence. Subject *A*'s results do correlate positively, but subject *B*'s correlate negatively. On the other hand the correlation between the M. S. for the right hand and degree of confidence is high, being .83. The correlation between the M. S. for the other two forms of presentation, simultaneous and successive, and the degree of confidence, however, is only .40. As one would suppose, we do find a certain correlation between lack of confidence and the number of doubtful judgments. Subjects *B* and *F* have the most doubtful judgments. They also are low in confidence. On the other hand, because a subject lacks confidence is no indication that he will have many doubtful judgments, as is shown in the case of subject *C*, who has only ten doubtful judgments and only thirty-seven *a* judgments out of seven hundred and fifty.

It will be recalled that the subjects were instructed that when they could not answer 'larger,' or 'smaller,' or 'equal,' they were to record the fact by naming it a *d* judgment and

then guessing either larger or smaller. There was found to be among all the judgments made in the half-year, including the preliminary trials of the main series, two hundred and twenty-nine such guesses when there was an objective difference. Of these, one hundred and sixty, or 70 per cent., were correct, and sixty-nine, or 30 per cent., incorrect. There were forty-nine guess judgments with objective equality, namely, when $d = 0$. Here, in contrast with the above distribution, there are almost an equal number of larger and smaller judgments, namely, twenty-six and twenty-three. These results conform to the laws of chance, although there are too few cases to be at all convincing, except in connection with the other results. Most of the guess judgments were divided between subjects *B* and *F*. Subject *B* had 72 per cent. correct and 28 per cent. incorrect; subject *F*, 71 per cent. correct and 29 per cent. incorrect. All the others showed the same tendency, but the numbers were too small to have any value except to strengthen the evidence of a predominance of correct answers by leaning in the same direction. It is clear, therefore, that although the subjective state of the subject was one of extreme doubt with no conscious attitude to swing him one way or the other, if he made a judgment without deliberation or hesitancy, he would have had many more correct answers than incorrect.

In thus making 'greater' or 'less' judgments, the subject probably assumes an attitude more closely resembling his set toward the other judgments. The 'doubtful set' is undoubtedly a change of attitude toward the experiment and should not be encouraged. It seems as if the instruction to guess 'larger' or 'smaller' gradually brings the subject into a constant attitude or to one approximating it. This seems to be borne out by the decrease in doubtful judgments. George says that ". . . the observers tended in the effort to maintain a constant receptive attitude to make passive judgments." Paradoxical as it may seem, the attitude when one guesses in the above situation is more nearly like the passive attitude as contrasted with the reflective, deliberate attitude of the doubtful judgment. George says further

that "the indication is that the way to maintain a constant attitude is to dispose oneself to make passive immediate judgments." [1] Our experience seems to support this. One's attention is diffused, as it were, over the experience, rather than concentrated now on this, now on that, criterion, and the response comes, it often seems, almost involuntarily. One is often surprised that the judgment is correct. There seems to be little doubt that the subjects of the psychophysical experiment should be trained in this attitude.[2]

In this connection, subject *B*'s two sets of results are of value. The ones in brackets, in which the doubtful judgments have been treated as equal judgments, give higher M. S.'s and lower precision than the results in which these doubtful judgments have been ruled out and judgment of *a*, *b*, or *c* certainty substituted. From the relation of the right to wrong judgments for this set, it is evident that if, instead of ruling out the 'guess' judgments and substituting others, the 'larger' and 'smaller' guess judgments had been used, the results would have been similar to those obtained from the former procedure and which have been used in the table unbracketed.

R. S. Woodworth recently reopened the question of the use of the equal and doubtful category.[3] He says on page 66 that "the main point is, of course, that the middle class of judgments, consisting of 'equal' and 'doubtful' judgments, is only a subjective definition, varying from individual to

[1] *Op. cit.*, p. 33.

[2] Several investigators have analyzed guess judgments and our results are in agreement with theirs. Pearce and Jastrow, On Small Differences of Sensation, *Memoirs of the National Academy of Sciences*, Vol. 3, 1884, found 60 per cent. correct guesses. Fullerton and Cattell found 60 per cent. and 65 per cent. correct judgments for two subjects. *Op. cit.*, p. 132. Whipple says, *op. cit.*, p. 83: "At the same time we found the subject had overlooked this element of his field of sensation, although his attention was directed with a certain strength toward it, so we marked his confidence as 0. This happened in cases where the judgments were so much effected by the difference of pressure as to be correct three times out of five." Griffing, *op. cit.*, p. 46, found the average of all correct guesses was 59 per cent. "In the case of some observers whose confidence was small, the percentage ran as high as 65 per cent. and 70 per cent."

[3] Professor Cattell's Psychophysical Contributions, *Arch. of Psychol.*, 1914, 30, 60–74.

individual. One individual insisting on a rather high degree of clearness or confidence for his 'greater' or 'less' judgments will throw a large portion of the cases into the middle class, while another individual will follow slighter indications and throw similar cases into the plus or minus column. The individuals differ in their standards of confidence, but not necessarily in their power to perceive a difference. Now since it has been proved that the division of the middle class between the other two classes results in a good balance of correctness, the individual who takes the more chances or has the lower standard of clearness will get the better record." Woodworth therefore continues to advocate forcing the subject into the false position of inhibiting equal and doubtful judgments in order to make results of all subjects comparable. This end can probably be attained better by the method outlined in this paper. In regard to the equal judgments, as was said before, none of the subjects seem to have had difficulty in getting the positive perception of equality (except the one subject with the preëstablished set, and even he made many such equal judgments). The doubtful judgments are allowed, but not included in the results.[1] The results showed no indication of avoiding a positive decision or of lapsing into a state of mild 'aboulia,' as Warner Brown intimated would happen under such conditions.[2] If the doubtful judgments were allowed to remain with the equal judgments in the results, Woodworth would be justified in saying that "the boundaries of the middle class are fixed by each subject without objective control and almost invariably differ from subject to subject, etc.," [3] for we saw how many categories the *d* judgments included. By counting only the very definite category of 'equality' in the sense of a decided impression of equality the boundary is very firmly fixed. By allowing the subjects who so desire to make doubtful

[1] G. W. Whipple repeated the doubtful tests until the observer made satisfactory judgments, but the number and distribution of these cases were recorded. *Op. cit.*, p. 417.

[2] The Judgment of Differences, *Univ. of California Publications in Psychology*, 1910, 1, 32–33.

[3] *Op. cit.*, p. 65.

judgments we gain two points. We keep the boundary sharply defined in his consciousness and by noting such judgments are able to make and allow for a change of attitude. These doubtful judgments tend to grow less with practice and would probably disappear in the latter half of a series of one hundred judgments. In other words, the method of impressing upon the subject the difference between judgments based on the actual perception of equality and those called equal because there was no judgment one way or the other probably had the beneficial effect of preventing too many equality judgments, and the knowledge that guess judgments would not stand as such kept the subject keyed to the proper amount of attention and probably prevented him from shifting his attitude.

Summary

1. The measurement of sensitivity for finger span with a standard of 5 cm. for the two most reliable subjects was .45 mm., or about 1 per cent., when one hand was used. It was in most cases higher when the stimulus was presented to the two hands, whether successively or simultaneously, but for one of the two most reliable subjects it changed very little. His M. S. remained relatively low for all three forms of presentation.

2. There was no decided agreement between subjective confidence and reliability. The best subject did have a high degree of confidence, but the second best had a low degree.

3. Decidedly more right judgments had a higher degree of confidence than wrong judgments.

4. When the subjects were forced to make judgments which had the quality of guesses, there were many more right than wrong answers.

A NEW OLFACTOMETRIC TECHNIQUE AND SOME RESULTS

BY HERBERT WOODROW AND BENJAMIN KARPMAN

University of Minnesota

The technique here to be described consists mainly in a new odorimetric method, which enables one to control accurately and easily the physical intensity of odoriferous stimuli. Besides a new odorimetric method, however, we have developed, by combining this method with the measurement of adaptation times, a new olfactometric technique, which should be of general usefulness in the study of the sense of smell. We shall first describe the odorimetric method, and explain its theoretical basis, and then give some results on adaptation time, showing in particular how adaptation time varies with the physical intensity of the odor.

Odorimetric Method.[1]—The fundamental principle of the odorimetric method is simply that the number of molecules given off to the air by any substance increases with its temperature. The number of odorous molecules in a given volume of atmosphere is here regarded as the physical intensity of the odor. At a given temperature, the number of molecules given off by an odorous substance in a closed space depends upon what is known as the vapor tension of the substance. The vapor tension increases with the temperature. Determinations of the vapor tensions for a wide range of temperatures have already been made by chemists in the case of a large number of odorous substances; so that about all the experimenter has to do is to regulate the temperature of the odor with which he wishes to experiment.

Briefly stated, the procedure is as follows: Some of the

[1] We are pleased to acknowledge indebtedness to Professor Hunter, of the chemistry department of the University of Minnesota, and to Professor Gibson, of the University of the Phillipines, for important suggestions, both as to method and apparatus.

odorous substance is placed at the bottom of each of a pair of Emmerling absorption tubes filled with glass beads. A moderate current of air, supplied by a water air-pump, is passed through these tubes and on into the nose of the subject. The temperature of the tubes containing the odorous substance is regulated by means of an Ostwald thermostat and accessories. In order to avoid the danger of precipitation during the transit of the odorous vapor from the Emmerling tubes to the nose, and also in order to avoid the unpleasantness which most odors produce when they become very strong, temperatures above 33° C. are not used. No other air is allowed to enter the subject's nose than that coming through the tubes. The subject breathes gently and regularly through his mouth, which is kept wide open. The air, in passing through the Emmerling tubes, becomes as saturated with the odorous vapor as the temperature permits. Just after leaving the Emmerling tubes, and just before entering the nose, the air is warmed, by being passed through a hot water bath, to a temperature of 33° C.; so that no matter what the temperature of the odorous substance in the Emmerling tubes, the mixture of air and odorous vapor entering the nose always has the same temperature.

Now, according to Avogadro's rule, equal volumes of gas or gas mixture, at the same temperature and pressure, have equal numbers of molecules. In all cases, then, that is, no matter what the substance used nor its temperature in the Emmerling tubes, the same number of molecules of gas mixture enter the nose in a given unit of time; for in all cases the air pressure and the temperature at the nose are constant. The percentage of the total number of molecules entering the nose composed by the odorous ones will be directly proportional to the vapor tension of the odorous substance in the tubes, no matter what the substance. The rise in temperature of the gases just before they enter the nose does not change the *proportion* of odorous molecules. Since, then, equal numbers of molecules of the gas mixture (air and odorous vapor) are entering the nose in each unit of time, and since the proportion of odorous molecules in this mixture is

directly proportional to the vapor tension of the odorous substance in the tubes, the absolute number of odorous molecules entering the nose in unit time is directly proportional to the vapor tension. Consequently, we may use as our units of intensity the vapor tension (expressed in millimeters of mercury). When the vapor tension doubles, the intensity doubles, that is, twice the number of odorous molecules enter the nose in given time. Or, from two different odorous substances, we will have a given unit of time the same number of molecules entering the nose, providing we keep these substances at such temperatures that their vapor tension is the same.

Since the phenomenon of vapor tension forms the basis of the preceding method, the theory of the method can be made clear only by a brief exposition of the concept of vapor tension. According to the kinetic theory of gases, the molecules of a gas are practically independent of each other and are briskly moving in all directions in straight lines. Because of the freedom of its molecules, a gaseous substance will expand, and fill any space presented to it. As it fills the vessel, the impact of the gas molecules on the walls of the vessel produces a certain pressure, which is the gaseous pressure. In the case of a liquid, the molecules still possess sufficient independence to slide over one another. In spite of the clinging together of the molecules of the liquid, some of them have sufficient motion to free themselves from their neighbors, and, leaving the liquid surface, become gaseous. The same is true, though to a lesser degree, of solids. Now, since the velocity of the molecules depends directly on the temperature, it is evident that as the temperature rises the number leaving the liquid will rise. These molecules which have left the liquid give rise to a pressure, which is called the vapor pressure. Hence the vapor pressure is a measure of the number of molecules of the liquid present above it, in a confined space.

If an odorous liquid be placed at the bottom of an enclosed vessel, the pressure of the molecules given off by it will then be exerted in all directions, against the liquid as well as in

other directions. After a certain time, owing to the vapor pressure, as many molecules will press back into the liquid as succeed in freeing themselves from it. The liquid and the vapor are then in equilibrium. Consequently, it may be said that the vapor tension of the liquid equals the vapor pressure.

Now, the number of molecules of a gas in a gas mixture is proportional to its partial pressure. In the gas mixture consisting of air and an odorous vapor, after equilibrium has been reached between the odorous vapor and its liquid, the partial pressure of the odorous vapor is equal to the vapor pressure, which, in turn, is equal to the vapor tension of the liquid at the given temperature. Hence the number of molecules of the odorous vapor in a mixture of air and vapor, in a confined space above the odorous substance, is proportional to the vapor tension of the odorous substance.

If two odorous vapors have the same partial pressure at the same temperature, then their molecules will be present in the mixture with air in equal numbers and in the same proportion or molecular concentration. (The same weights of substance will not be present, unless their molecular weights are the same.) If two odorous vapors have the same partial pressure, but at different temperatures, then they will not have equal numbers of molecules. They will, however, have the same molecular concentration, that is, the same proportion of the total number of molecules of the mixture of odorous vapor and air; and they may be made to have the same absolute number of molecules by equating the temperatures of the gas mixtures. In a gas mixture, the molecular concentration of the components remains the same with rise in temperature. If the temperature of a mixture be lowered, liquid will settle out, the amount depending, for a given change, on the shape of the vapor tension curve (vapor tension plotted against temperature). It is necessary, therefore, for our purpose, to equate the temperatures of our mixtures of air and odorous vapor by *raising* the temperature, at least that of the mixture with the lowest temperature. Otherwise, we would incur an error due to precipitation.

Actually, as already stated, we *raise* the temperature of all mixtures to 33° C., before they enter the nose.

Lastly, it should be kept in mind that the molecular weight of a substance has no effect on the number of molecules which are necessary to produce a given pressure. There are the same number of hydrogen molecules present in a unit volume of pure gas at a given temperature and pressure as there would be of any gas, however heavy. This is simply a restatement of Avogadro's rule.

The apparatus consists of well-known, inexpensive chemical pieces. It is made up, first, of a mechanism for injecting a steady current of pure air through the Emmerling absorption tubes; second, of the Emmerling tubes, which contain the odorous substance; third, of mechanism for controlling the temperature of the tubes; and fourth, of two glass tubes which conduct the odorous air from the Emmerling tubes through a hot water bath into the nose. The tubing used consisted exclusively of glass. The necessary junctions were effected by especially devised aluminum joints. These were made of two short aluminum tubes, each of which fitted over one of the two glass tubes it was desired to hold together. The two aluminum tubes were threaded so that one screwed into the other. Disconnecting the tubes was therefore almost as easy as unscrewing the tip of a fountain pen. The short aluminum tubes were attached, at the end away from the joint, to their glass tubes either by means of a paste of lead oxide or by means of adhesive tape, thus avoiding the use of rubber.

The air current was supplied by means of a water air-pump or waterblast connected with the laboratory hydrant. Before entering the Emmerling tubes, the air passed through two Novy jars. The first, containing a solution of potassium permanganate, served to rid the air of impurities and the second, containing calcium chloride, to rid it of moisture. Upon smelling the air as it came from the second Novy jar, no sign of an odor was perceptible.

The air next passed through the Emmerling absorption tubes, where it became saturated with the odorous vapor.

We began by using three Emmerling tubes, the first placed outside of the thermostat and the last two inside it. By preliminary trials, however, with propyl alcohol, it was found that the adaptation time was, on the average, very nearly the same whether we used two tubes or three. We consequently concluded that the use of two tubes was sufficient to bring about saturation, and thereupon discarded the tube outside of the thermostat.

The Emmerling tubes were placed in a large Ostwald thermostat, with motor-driven fan. The tank of the thermostat was filled either with ice and water or water only. The temperature was regulated automatically by means of a Reichert-Novy gas regulator and a Bunsen burner. One could probably get along with an ordinary tub and a supply of ice and hot and cold water.

After leaving the second Emmerling tube, the odorous air current passed immediately into a tube about one yard long but bent so as to fit into a pan of water about ten inches long. The water in this pan was kept at 33° C., by means of a Bunsen burner, regulated, when necessary, by hand. In the work with propyl alcohol, in which the thermostat temperatures used were all below room temperature, this hot water bath was omitted. Lastly, the air passed into a short tube, almost vertical, which entered the subject's nostril. The odorous air issued from this tube into the nose at a rate of 21.8 c.c. per sec. The internal diameter of the tube was 5.0 mm., and its thickness 1.0 mm. The monorhinic, or one-nostril, method, with the other nostril held closed by the finger, was preferred to the dirhinic, or two-nostril, because with it, the stream of air has a greater force, thus reducing the possibility of interference from the breathing (through the mouth) and assuring a continuous stimulation, during expiration as well as inspiration. Of course a stronger air current than the one we used could be employed, but it might have to be passed through a larger number of Emmerling tubes to assure complete saturation. The odorous particles blown into the nose probably are not carried by the air current directly into contact with the *regio olfactoria*. They are

carried close to it and travel the rest of the distance, an exceedingly short one, by diffusion. The diffusion rates of different odorous vapors, while not the same, would not be appreciably different for such a short distance.

The accompanying diagram shows the various parts of the apparatus and the way they were connected together. In this diagram, all parts of the thermostat, except the immersion tank, have been omitted.

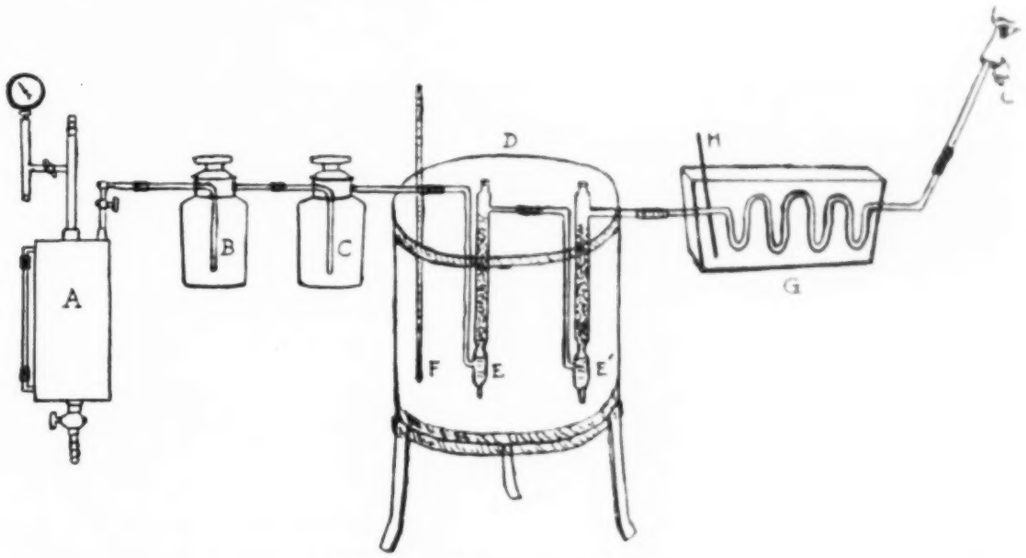

FIG. 1. Diagram of Apparatus.

A, air-blast.
B, Novy-jar, partially filled with solution of potassium permanganate.
C, Novy-jar, partially filled with calcium chloride.
D, thermostat tank.
E, *E'*, Emmerling absorption tubes filled with glass beads. (The odorous substance is placed at the bottom of these tubes.)
F, thermometer.
G, hot water bath.
H, thermometer.

The preceding odorimetric method is the only method so far described which enables one to stimulate the nose continuously for any desired time by means of an air current containing a known proportion of odorous molecules. It measures physical intensity of odor by means of units which are the same for different odorous substances, and this intensity does not vary with the duration of the stimulation.

Without going into a detailed criticism of the numerous

odorimetric methods employed in the past, we should like to point out one source of error, in addition to those commonly noted,[1] that appears to have been overlooked. This consists in an ignoring of the facts of vapor tension. For example, in the case of solutions, even if the solvent is odorless, the intensity of the odorous vapor at a given temperature does not vary, as often assumed, directly with the concentration of the solution. The vapor tension (or partial pressure) of ammonia in solution of water, at 19.9° C., is, for a 10.15 per cent. solution by weight, 80.6 mm. of Hg, while for a 23.37 per cent. solution, it is 302.4 mm. of Hg.[2] Thus, while the concentration of the solution increases 130 per cent., the physical intensity of the odorous vapor increases 275 per cent. In other cases, the increase in the intensity of the odor is less than proportional to the increase in the concentration of the solution.

Complications arise also with solids, as used, for example, in Zwaardemaker's method, a method which, in spite of the fact that it is known to suffer from serious errors,[3] appears to enjoy some popularity. It is based on the assumption that the mass of odorous particles given off from a solid body or the surface of a liquid, is, *other things equal*, proportional to the time and to the area of the surface exposed. This is true of a given substance in an unclosed or very large space. But in the application of this principle to an odorimetric method, Zwaardemaker makes use of an air current produced by inhalation through a tube composed of the odorous substance. This air current passes over the sections of the exposed surface in succession, and, as it does so, presumably becomes more saturated. At a given temperature, however, its saturation is limited. If the air, in passing over the first part of the surface, becomes saturated, its passing over

[1] For a history of previous odorimetric methods, see Gamble, *Amer. J. of Psychol.*, 1899, p. 98 ff. Among the common sources of error we may note the following: Adhesion of odorous particles to the walls of tubes or vessels; the use of odorous solvents; variation with the mode of breathing, and, in the case of bottled gases, with the time the bottle has been unstoppered; lack of control of the factor of diffusion (in Frölich's method); and incomparability of the intensities of different odors.

[2] Landolt and Börnstein, *Physicalische Chemische Tabellen*, 4th ed., 1912, p. 429.

[3] See Gamble *op,. cit.*, pp. 101–104.

additional surface will result in no increase in the intensity of the odor. Even though, as is probable, the air current in passing over the first part of the surface does not become saturated, its saturation in passing over a second portion of equal extent need not double, nor in passing over a third and fourth portion, triple and quadruple. Thus Zwaardemaker's method, in common with other methods, is theoretically unsound, because of neglect of the facts of vapor tension. The theory of vapor tension is fundamental to the understanding of the genesis of odors and must be taken into careful consideration by any odorimetric method that derives the odors from solids or liquids.

Adaptation Time and Intensity.—Having devised a satisfactory method for the control of the physical intensity of odor, we proceeded to study how the time required for adaption varies with the intensity. The first substance tried was propyl alcohol. With this odor four subjects were used. Later, experiments were made with two more substances, camphor and naphthalene, using three subjects. None of the subjects were smokers.

The vapor tensions of these substances, at the temperature used, is given in the adjoining table. Under the head of intensity, we have taken as units different vapor tensions in different cases. With propyl alcohol, we began with a temperature of 10° C., which gives a vapor tension of 7.4 mm. We arbitrarily called this two units of intensity, and then arranged our other temperatures so as to give a range of intensities from one to four units. With naphthalene and camphor we took as our unit of intensity a vapor tension of .05 mm. and adjusted our temperatures to give a range of from one to five units with naphthalene and from two to six with camphor.[1] The determinations of adaptation time were made in such order that the various intensities were mixed irregularly.

By way of practice, preliminary trials were made with

[1] The vapor tensions were taken from Landolt and Börnstein's *Physicalische Chemische Tabellen*, 4th ed., 1912. Inasmuch as the temperature is usually given in these tables by steps of five degrees (0, 5, 10, etc.), it was often necessary to interpolate in order to obtain the temperature required for the desired vapor tension.

each subject with each odor. The number of these preliminary trials was particularly large in the case of propyl alcohol, the first odor used. Propyl alcohol differed from the other odors in that with it there occurred a progressive shortening in the adaptation time with practice. With propyl alcohol, all trials were regarded as preliminary until this progressive shortening seemed to have ceased. In some cases no shortening was apparent after the second trial. In other cases it continued through five trials. With the other odors, where shortening with practice was not evident, each subject was given one preliminary trial with each intensity.

TABLE I.

SHOWING VARIATION OF VAPOR TENSIONS WITH TEMPERATURE

Substance	Temperature	Vapor Tension	Intensity
Propyl Alcohol	0.6° C.	3.7 mm. Hg.	1x
	10.0 "	7.4 " "	2x
	15.8 "	11.1 " "	3x
	19.7 "	14.8 " "	4x
Napthalene	11.0 "	.05 " "	1x
	24.5 "	.10 " "	2x
	31.0 "	.15 " "	3x
	34.4 "	.20 " "	4x
	36.9 "	.25 " "	5x
Camphor	10.0 "	.10 " "	2x
	20.0 "	.15 " "	3x
	25.8 "	.20 " "	4x
	29.5 "	.25 " "	5x
	31.5 "	.30 " "	6x

The preliminary results with propyl alcohol show that in some cases a certain amount of practice may be necessary before the subjects become reliable. They may have to learn to neglect many irrelevant sensations: taste, touch, burning, etc. With propyl alcohol, which, particularly at the higher intensities, produces a burning sensation, subjects gave the introspection that there was some difficulty in distinguishing between smell and touch and smell and taste. As the work continued, it became easier to discriminate the smell from the other sensations.

All the subjects reported that they could smell the odor

continuously, that is, during expiration as well as inspiration. This report seems natural enough, as the air current carrying the odorous vapor entered the nose at a steady rate, 21.8 c.c. per sec., while the subject breathed gently and exclusively through his wide-open mouth. One of the subjects, however, found that as the odor weakened with the progress of adaptation, it became much harder to be sure of it during expiration than during inspiration. All the subjects found it difficult

TABLE II.

ADAPTATION TIMES WITH PROPYL ALCOHOL

Intensity		1 *x*	2 *x*	3 *x*	4 *x*
Temperature		0.6°C	10.0°C.	15.8°C.	19.7°C.
Vapor Tension		3.7 mm.	7.4 mm.	11.1 mm.	14.8 mm.
Subj. Kn	Trial 1	102	144	180	240
" "	" 2	64	114	126	210
" "	" 3	110	96	115	240
" "	" 4	72	130	175	210
" "	" 5	55	144	156	120
" "	" 6	82	74	135	195
" "	Average	**81**	**117**	**148**	**203**
" "	*M. V.*	*17*	*23*	*22*	*29*
" Dy	Trial 1	24	126	135	255
" "	" 2	72	75	180	375
" "	" 3	78	126	130	105
" "	" 4	65	85	135	165
" "	" 5	51	96	150	210
" "	" 6	54	105	138	225
" "	Average	**57**	**102**	**145**	**223**
" "	*M. V.*	*15*	*17*	*13*	*62*
" Dl	Trial 1	115	85	155	285
" "	" 2	60	115	162	245
" "	" 3	45	90	192	105
" "	" 4	85	85	184	190
" "	" 5	186	130	156	145
" "	" 6	75	163	141	240
" "	Average	**94**	**111**	**165**	**201**
" "	*M. V.*	*37*	*28*	*15*	*56*
" Pn	Trial 1	72	95	175	270
" "	" 2	60	125	120	160
" "	" 3	90	85	100	180
" "	" 4	55	95	165	180
" "	" 5	70	95	150	135
" "	" 6	52	125	105	165
" "	Average	**67**	**103**	**136**	**182**
" "	*M. V.*	*10*	*15*	*27*	*29*
All Subjects	Average	**75**	**108**	**148**	**202**
" "	*Ave. M. V.*	*19*	*21*	*19*	*44*

to tell exactly when the odor was no longer present. They were instructed to say 'Now,' just as soon as they could decide for the first time that they were getting no sensation of odor. These instructions were due to the fact that in some of the preliminary trials we obtained statements as follows: "Twice it seemed as if the smell disappeared, only the burning sensation remaining. Then the smell appeared a third time before vanishing for good."

Table III.

Adaptation Times with Napthalene

Intensity		1.x	2.x	3.x	4.x	5.x
Temperature		11.0° C.	24.5° C.	31.0° C.	34.4° C.	36.9° C.
Vapor Tension		.05 mm.	.10 mm.	.15 mm.	.20 mm.	.25 mm.
Subj. Ud	Trial 1	67	87	97	104	147
" "	" 2	73	84	94	98	129
" "	" 3	80	61	102	113	126
" "	" 4	65	79	86	98	131
" "	" 5	71	82	96	96	128
" "	" 6	76	77	98	114	136
" "	Average	**72**	**78**	**96**	**104**	**133**
" "	*M. V.*	*4*	*6*	*4*	*7*	*7*
" Fe	Trial 1	66	79	77	116	111
" "	" 2	84	74	88	94	131
" "	" 3	71	73	103	99	119
" "	" 4	72	69	96	103	129
" "	" 5	58	76	84	98	109
" "	" 6	53	73	89	104	124
" "	Average	**67**	**74**	**90**	**102**	**121**
" "	*M. V.*	*8*	*3*	*7*	*5*	*8*
" Kn	Trial 1	45	68	76	79	108
" "	" 2	50	72	86	93	103
" "	" 3	60	53	71	84	102
" "	" 4	49	67	73	69	101
" "	" 5	52	62	64	81	86
" "	" 6	48	62	74	77	100
" "	Average	**51**	**64**	**74**	**81**	**100**
" "	*M. V.*	*4*	*5*	*4*	*6*	*5*
All Subjects	Average	**63**	**72**	**87**	**96**	**118**
" "	*Ave. M. V.*	*5*	*5*	*5*	*6*	*7*

In the accompanying tables, Tables II.–IV., we give the adaptation time obtained at each trial. Six trials were made with each subject at each of the intensities. Each of the tables shows, for one of the three odors used, how adaptation

time varies with physical intensity. The absolute measure of the physical intensity in each case is the vapor tension, which is given at the heads of the columns in terms of millimeters of mercury. These vapor tensions may be taken as strictly comparable absolute measures of intensity, no matter what the odor. Figs. 2 and 3 show the variation in the average adaptation time with the physical intensity of the odors.

TABLE IV.

ADAPTATION TIMES WITH CAMPHOR

Intensity		2.r.	3.r.	4.r.	5.r.	6.r.
Temperature		10.0° C.	20.0° C.	25.8° C.	29.5° C.	31.5° C.
Vapor Tension		.10 mm.	.15 mm.	.20 mm.	.25 mm.	.30 mm.
Subj. Ud	Trial 1	102	160	184	231	265
" "	" 2	97	133	175	216	252
" "	" 3	98	142	186	193	236
" "	" 4	104	152	169	208	228
" "	" 5	84	134	142	222	202
" "	" 6	82	132	145	230	215
" "	Average	**95**	**142**	**167**	**217**	**233**
" "	*M. V.*	*8*	9	*15*	*11*	*18*
" Fe	Trial 1	72	127	144	179	204
" "	" 2	67	106	138	166	200
" "	" 3	71	84	126	158	184
" "	" 4	65	91	118	154	192
" "	" 5	69	108	136	166	179
" "	" 6	65	96	134	157	187
" "	Average	**68**	**102**	**133**	**163**	**191**
" "	*M. V.*	*3*	*12*	*7*	*7*	*8*
" Kn	Trial 1	63	101	124	151	198
" "	" 2	60	97	112	140	189
" "	" 3	58	106	115	158	204
" "	" 4	59	85	121	134	180
" "	" 5	56	96	109	140	200
" "	" 6	58	97	115	135	196
" "	Average	**59**	**97**	**116**	**143**	**193**
" "	*M. V.*	*2*	*4*	*4*	*8*	*7*
All Subjects	Average	**74**	**114**	**139**	**174**	**206**
" "	*Ave M. V.*	*4*	*8*	9	9	*11*

The data lead to definite conclusions. They show a marked increase in adaptation time with increase in the physical intensity of the stimulus. It is true that the variation in the adaptation time of any individual for any particular

intensity of odor is rather large. It is not so large, however, but that averages based on six determinations give fairly regular curves. This is seen by the fact that if the curves

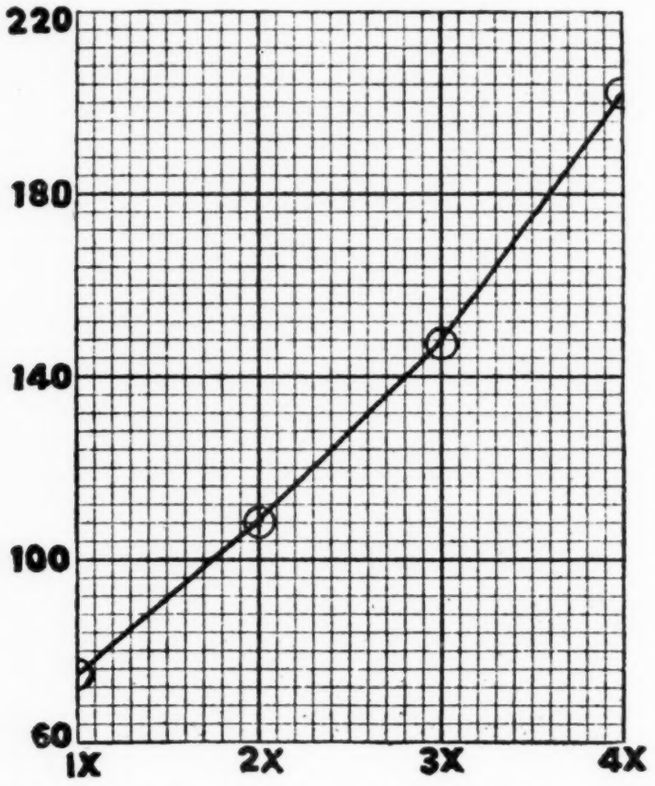

FIG. 2. Variation in Adaptation Time with Physical Intensity, Propyl Alcohol.

are plotted for the individual subjects, they will be found to have the same general slope. This fact justifies our averaging the results obtained with all subjects and plotting the curves of the averages, in order to determine the general relationship between adaptation time and physical intensity.

A study of these curves (Figs. 2 and 3) shows that they are fairly regular. They are at least regular enough to make it certain that adaptation time is a function of physical intensity and that this function can be determined with a high degree of precision. The data here presented are probably not sufficient in number nor in range to justify a final decision as to the mathematical law which prevails. It will be seen, however, that all three of the curves approximate straight lines, the formula for which is very simple. It is

that $t = K + k \cdot I$, in which t represents the adaptation time in seconds, K and k determinable constants and I the physical intensity. We may say, then, that adaptation time varies

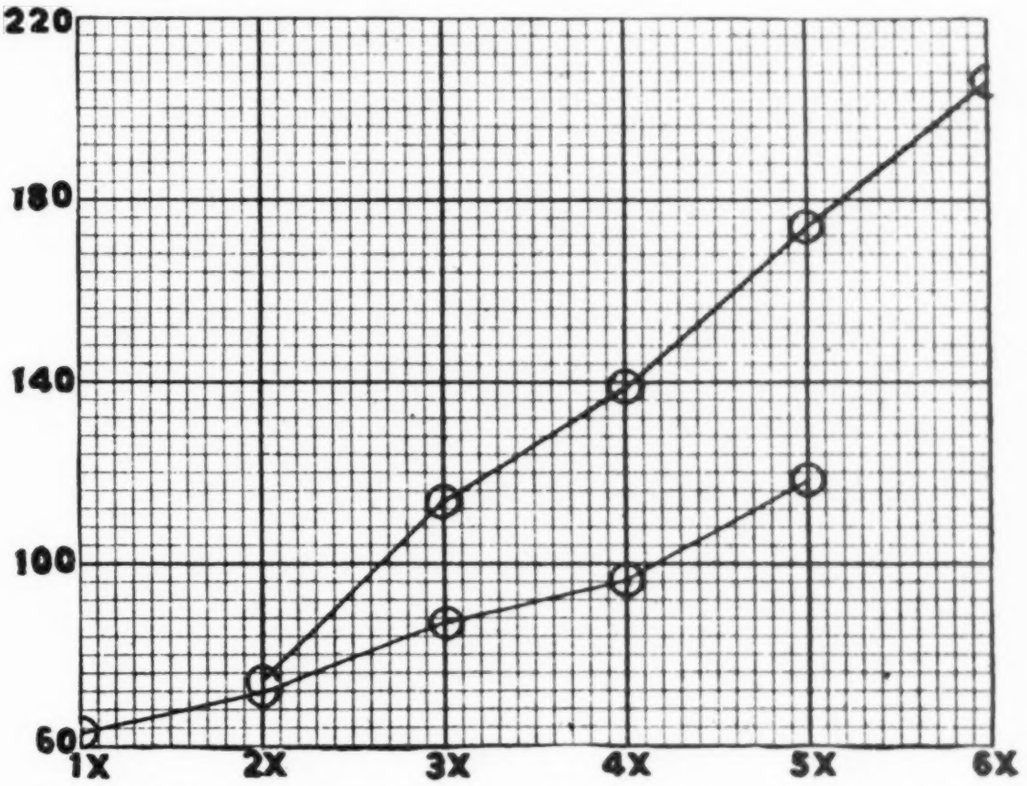

FIG. 3. Variation in Adaptation Time with Physical Intensity. Upper Curve, Camphor; Lower Curve, Naphthalene.

as a constant plus the product of another constant into the intensity of the stimulus; or, more simply, that *increasing the physical intensity of an odor by equal steps causes an increase in adaptation time by equal steps.* The *increase* in the one is directly proportional to the *increase* in the other. The very nature of this relationship suggests that it holds only within limits, at least that there is a lower limit, probably somewhere near the sensation threshold.

The following table shows the deviations of the adaptation times obtained experimentally from the most probable true ones on the assumption that $t = K + k \cdot I$. When the large mean variation and the small number of the individual determinations is kept in mind, it will be seen that the deviations are not serious.

TABLE V.

A COMPARISON OF OBTAINED ADAPTATION TIMES WITH THOSE CALCULATED BY THE LAW $t = K + k \cdot I$

Odor	Intensity	Obtained t	Calculated t	d
Propyl Alcohol.....	3.7 (1 x)	75	70	+ 5
$K = 28$	7.4 (2 x)	108	112	− 4
$k = 42$	11.1 (3 x)	148	154	− 6
	14.8 (4 x)	202	196	+ 6
Naphthalene.....	.05(1 x)	63	61	+ 2
	.10(2 x)	72	74	− 2
$K = 48$	.15(3 x)	87	87	0
$k = 13$	.20(4 x)	96	100	− 4
	.25(5 x)	118	113	+ 5
Camphor.....	.10(2 x)	74	76.5	− 2.5
	.15(3 x)	114	109.0	+ 5.0
$K = 11.5$	.20(4 x)	139	141.5	− 2.5
$k = 32.5$	.25(5 x)	174	174.0	0
	.30(6 x)	206	206.5	− 0.5

Intensity of stimulus has hitherto been regarded largely as though its sole mental function were intensity of sensation. We have shown that the duration of the sensation is also a function of intensity of stimulus, one which, on account of its easy measurement, as adaptation time, should be of fully as much interest as intensity of sensation in the study of the senses of smell, touch and temperature.

In conclusion, we desire to mention a few of the problems immediately suggested as suitable for study by the technique described in the preceding pages. One of these is the relation of adaptation time to intensity of sensation. Do two odors which have the same adaptation time have also the same intensity of sensation? We have not determined the answer to this question, but, on the basis of what observations we have made, we would incline towards an affirmative. At any rate, adaptation times should serve as valuable indirect measurements of sensational intensity.[1] Another interesting

[1] The relationship between the intensity of stimulus and the intensity of sensation, in the case of smell, has been studied by Gamble (The Applicability of Weber's Law to Smell, *Amer. J. of Psychol.*, 1899, 82–143). The interrelationship of the three factors, physical intensity, sensation intensity and adaptation time should be worked out not only for the sense of smell, but also for the temperature senses and the sense of touch.

question is the relationship of adaptation time to the classification of odors. Could we not accurately rank the qualitative resemblance of any number of odors to a standard by ranking them in the order in which their adaptation time is decreased by adaptation to the standard? For such work, it might be desirable first to equate the adaptation times of the different substances. This could be done by the present technique, aided by calculation. Using this technique, then, it might very well be possible to explore the odor-prism of Henning[1] and determine the correctness of its various dimensions. These problems are sufficient to serve as illustrations of the value of a sound odorimetric technique and of the usefulness which such a technique may acquire when combined with the measurement of adaptation times.

[1] Henning, Der Geruch, *Zsch. f. Psychol. u. Physiol. d. Sinnes.*, 1915, p. 254.

THE MEMORY VALUE OF MIXED SIZES OF ADVERTISEMENTS

BY HENRY F. ADAMS

University of Michigan

When advertisements are used in memory experiments, difficulties are encountered which are easily avoided by the use of more homogeneous materials. In the first place, the purpose of the advertising test is somewhat different from that of the ordinary memory experiment, for with advertisements, the desired result is the recall either of the name of the commodity or a description sufficient to identify it, or the recollection of some of its attractive qualities. Should any of these be remembered, the advertisement receives full credit.

Secondly, advertisements differ from each other in every conceivable way both on the objective and the subjective sides. The greater number of experiments which have been performed have dealt with the objective factors of size and frequency, undoubtedly because they are easier to control. Since advertisements differ also from the subjective standpoint, it is necessary to take some account of it as well. This is usually done by having a sufficient number of advertisements of the same size or of the same number of presentations so that it seems possible to disregard other differences and assume that variations due to interest factors will be compensated for in the averages. That this assumption is not unwarranted is indicated by the close similarity in the results of different investigators.

It might appear from the practical uniformity of results that the effects of the objective factors had been sufficiently determined and that further experimentation in that direction would be superfluous. But new considerations are constantly appearing which suggest further experiments.

The purpose of this paper is to give the results of an

experiment in which different combinations of sizes of advertisements were given four presentations. In securing the material, four advertisements of the same commodity were selected, but not all four were of the same area. There might be three full pages and one half; or two full and two half; or one full, one half and two quarter pages of the same commodity. Necessarily the different advertisements of the same article were not identical in content, but exhibited a considerable range of variation. In all, each combination of size was represented four times so that the following conditions were met:

4 firms using 3 full and 1 half page,
4 firms using 3 full and 1 quarter page,
4 firms using 2 full and 2 half pages,
4 firms using 2 full and 2 quarter pages,
4 firms using 2 full, 1 half and 1 quarter page,
4 firms using 1 full, 2 half and 1 quarter page,
4 firms using 1 full, 1 half and 2 quarter pages,
4 firms using 1 full and 3 half pages,
4 firms using 1 full and 3 quarter pages,
4 firms using 3 half and 1 quarter page,
4 firms using 2 half and 2 quarter pages,
4 firms using 1 half and 3 quarter pages,
11 firms using 1 full page.

It will be seen that every possible combination of four presentations, using mixed sizes, was employed. These advertisements were pasted upon cardboard the size of the standard magazine page, no two advertisements of the same commodity appearing on the same or neighboring pages. This dummy was handed to the subject, who was asked to look it through in the same way that he would the advertising section of a magazine, except that he was to be sure that he had seen every advertisement. If one did not interest him, however, he was told to skip it and go on to the next. He was requested to spend about 20 minutes on the dummy, but might take more or less time if he desired. In this way, 35 subjects were tested.

Before giving the results, it is necessary to outline briefly the other experiments which have been performed to test the effect of size and frequency of insertion of advertisements upon their memory value. The reasons for this will appear shortly.

1. *The Effect of Size.*—This topic has been investigated by Scott,[1] Strong,[2] Starch,[3] and Adams.[4] While some of the tests were carried on by the recognition method, others by the recall, the form of the curve and the ratios were quite similar, though the absolute amount remembered was somewhat higher with the recognition test.

Another difference was the material used. Scott, Starch, and Strong in certain of his experiments, used magazines, which prohibited the selection of unfamiliar advertisements. The result was that certain of the full-page and a lesser number of half-page advertisements received undue credit because they were already well known to the reader. Starch devised an ingenious correction for familiarity and found that a full page should be allowed 63 per cent. of its obtained credit, a half page 73 per cent. and a quarter page 100 per cent., but it seems probable that he overcorrected, thus depriving the larger sizes of a certain amount of deserved credit.

A third difference was the time which intervened between presentations. The usual method has been to give the entire series in a period of less than an hour and to test immediately. Strong, however, presented 4 dummies at intervals of a few minutes, 1 day, 1 week and 1 month. In the first three cases, the test took place 4 weeks after the first set of advertisements had been seen, in the last, 16 weeks after the first dummy had been read through. If no elaborate corrections are made for probable differences in memory value owing to the irregular lapse of time, and this ought to be rectified in the ratios themselves, the following figures appear:

[1] Scott, W. D., 'The Psychology of Advertising,' 168–169.

[2] Strong, E. K., Jr., Psychol. Rev., Vol. 21, 137 ff. J. of Exper. Psy., Vol. 1, pp. 319–338.

[3] Starch, D., 'Advertising,' pp. 30 and 48.

[4] Adams, H. F., *J. of Phil., Psychol., etc.*, 1916, 13, 141–152. 'Advertising and Its Mental Laws,' pages 226–237.

Interval	Size of Advertisements		
	¼ Page	½ Page	1 Page
None	1.00	1.94	3.29
1 day	1.00	1.53	2.39
1 week	1.00	2.24	3.60
1 month	1.00	1.47	2.23

In the other experiments, all the advertisements were contained in one dummy, so the results are more nearly comparable with Strong's first set of ratios. The average of these other tests is given below:

Size of Advertisements		
¼ Page	½ Page	1 Page
1.00	1.87	3.16

These ratios are significantly similar to Strong's first set. However, if correction is made for familiarity in the manner suggested by Starch, these figures appear:

Size of Advertisements		
¼ Page	½ Page	1 Page
1.00	1.37	1.98

With such correction, the memory value of size varies approximately as the square root of the area; without correction, approximately as the 1.15 root of the area. The true figure probably lies somewhere between the two values. That both sets of figures should approximate so closely to simple root curves is striking. Strong's figures indicate that size is less and less effective as the repetitions occur further and further apart.

2. *A. The Effect of Frequency of Insertion. Duplicates Used.*—Less work has been done on frequency than on size. The general method employed has been to prepare a dummy, or series of them, in which quarter, half and full pages appear one, two, and four times. In part of the experiments the advertisements have been presented together; in others varying intervals have existed between the successive appearances. Strong investigated the latter condition in the experiment already mentioned under the topic of size. Working his

results through in ratios, as was done before, these figures are obtained:[1]

Interval	Number of Presentations		
	1	2	4
None	1.00	1.52	1.84
1 day	1.00	1.57	2.05
1 week	1.00	1.64	2.35
1 month	1.00	1.19	1.53

Adams, whose advertisements were all shown in one dummy, obtained these ratios for 1, 2 and 4 presentations: 1.00, 1.49, 2.57. If these are averaged with the corresponding figures of Strong's, the values immediately below are found:

Number of Presentations		
1	2	4
1.00	1.51	2.20

These ratios cannot properly be corrected for familiarity, since they were obtained by giving one insertion a credit of 1.00, no matter what its size might be, and putting the other frequencies in relative terms. They are between the corrected and uncorrected ratios for size and show that the effectiveness of frequency varies approximately as the 1.7 root of the number of insertions. From these figures alone it is impossible to determine which is more effective, size or repetition. However, from Strong's figures, quoted above, it appears that, on the average, size has greater effect than frequency of presentation. The same point is brought out by a consideration of results from certain mail-order campaigns.

With the exception of the one month interval, Strong's data indicate that the longer the time which elapses between successive insertions of duplicated advertisements, the more effective the second and fourth insertion will be. This is quite the opposite of the findings for size.

[1] In the tables adapted from those given by Strong, the unit of presentation or of size is given throughout the value of 1.00, for we were interested in obtaining the relative effectiveness of multiples of the units. The subject matter of this paper does not make desirable a discussion of the effect of the varying intervals upon the memory value of the unit.

2. *B. The Effect of Frequency of Insertion. Variations Used.*—It has been found that variation has a much greater effect upon memory than duplication. The experiment[1] yielded the following ratios:

Frequency of Presentation		
1	2	4
1.00	2.63	4.05

Since only 40 subjects were employed in this investigation, another similar experiment was performed with different materials in which 122 persons were tested. The second experiment gave these ratios:

Frequency of Presentation		
1	2	4
1.00	2.21	4.15

These two tests, performed on different groups of students, with different material and the experimental work done by different men, produced results which show the same general tendency. If a weighted average of the two is determined, in which the number of subjects used in each experiment is taken into account, the following ratios are obtained:

Number of Presentations		
1	2	4
1.00	2.33	4.13

From these experiments, it is justifiable to state that the effectiveness of variation varies almost, but slightly more than, directly with the number of presentations.

These are the more important experiments which have been performed upon the memory value of advertisements. Is it possible to take what they show, accept it provisionally, and deduce further laws? It will be interesting to try. In order to do so, however, certain definite assumptions must be made. These will now be considered.

It will be recalled in connection with the memory value of size that two sets of ratios were given, one uncorrected for familiarity, the other overcorrected, and the true value

[1] Adams, H. F., *J. of Phil., Psychol., &c.*, 1916, 13, 151–152.

was to be found between the two. Starch, who grants that his method was a makeshift, had his subjects write a list of all the commodities with which they were familiar, through advertising or any other means. The number of times each commodity was mentioned was then subtracted from the number of times it was remembered in the experiment and the remainder was supposed to be the true value of the advertisement. This surely is overcorrecting, for even advertisements which were wholly unfamiliar would have some memory value with these subjects and Starch's method overcorrects by at least that amount. Since Strong found that the average firm was remembered by about 50 per cent. of those experimented upon immediately after reading the dummy, it seems fair to split the difference between the results obtained with and without correction. This is done by giving a half-page advertisement 86.5 per cent. and a full page 81.5 per cent. of the obtained credits. When this is done, the following ratios are obtained for size:

Size of Advertisement		
¼ page	½ page	1 page
1.00	1.62	2.57

It must be remembered that the figures refer to those found when the presentations and the test all took place in a few minutes. These ratios vary approximately as the 1.45 root of the area. For the theoretical considerations which follow they will be taken as the working basis. There is justification for them because we are fairly sure, both from experimental evidence and from business returns, that doubling or quadrupling the size produces greater effect upon memory than a corresponding increase in the number of presentations with duplicated advertisements.

In the table immediately following is given a summary of the findings which will be used:

	Units of Stimulation			Efficiency
	1	2	4	
Size	1.00	1.62	2.57	1.45 root
Duplication	1.00	1.51	2.20	1.70 root
Variation	1.00	2.33	4.13	Almost directly

Since, in the experiment which serves as a basis for this paper, the advertisements of the same commodity were variations and also differed in size, it is possible to take the two tables dealing with these factors and put them together. It will then be found that certain very definite theoretical conclusions will appear. This table follows:

Size	Frequency of Presentation			
	1	2	3	4
¼	1.00	2.33	3.30	4.13
½	1.62	3.78	5.35	6.70
1	2.57	6.00	8.48	10.60

The figures are obtained by multiplying the ratios for size, in the perpendicular column under 1, by those for frequency, in the horizontal column opposite ¼. The values for frequency show the effect of 1, 2, 3 and 4 insertions of the advertisement. The effect of the second insertion may be determined by subtracting the ratio of one insertion from that of two; of the third, by subtracting that of two from that of three; and so on for the rest of the combinations. When this is done, these figures are obtained:

Size	Appearance Number			
	1	2	3	4
¼	1.00	1.33	0.97	0.83
½	1.62	2.16	1.57	1.35
1	2.57	3.43	2.48	2.12

If we wish to determine the probable memory value of a full page, followed by a half, followed by a quarter, we find in the table the effectiveness of the first presentation of the full page, 2.57, of the second presentation of the half page, 2.16, of the third presentation of the quarter page, 0.97, and add them together, getting 5.70. Similarly, if we wish to invert the order and determine the utility of a quarter, followed by a half, followed by a full page, we add the corresponding numbers in the table, 1.00, 2.16, 2.48, the sum being 5.64. This shows that the former order of presentation is slightly better than the latter.

But we have entirely disregarded one factor, and a very important one, too. It seems entirely plausible that a half page coming after a full page should be more effective than the same half page following a quarter page advertisement of the same commodity. This would be necessary if we take account of the conditions of attention, for we know in a general way that the direction of attention is determined by the relative vividness of the ideas, or by the relative permeability of brain tracts. A full page must, on the average, decrease the resistance of synapses more than a quarter page and leave the brain better prepared for attending a second time to the same or a similar object. Consequently, a half page coming after a larger size must have a higher value than one following a smaller size. The actual amount is conjectural in the present state of our knowledge. As a working basis, we have assumed that an advertisement preceded by one of double the area should receive an increase of 10 per cent. of its actual value, whereas, when the foregoing advertisement is half the area, it gets a corresponding reduction of 10 per cent. Likewise, when a quarter page followed a full, it received an arbitrary increase of 20 per cent. and a full page coming after a quarter, a decrease of 20 per cent. Similarly, when a full page was followed by 3 half pages, the first half page received an increase of 10 per cent., the second of 5 per cent. and the third of 2.5 per cent. These figures are intended to be merely illustrative. It is hoped that from a series of experiments now going on more definite and trustworthy values may be assigned. Some correction, however, is necessary.

If we do the examples given above, taking account of the corrections just outlined, the following is obtained:

Full page	2.570
Half page	2.16 + .216 = 2.376
Quarter page	0.97 + .097 = 1.067
Sum	6.013
Quarter page	1.000
Half page	2.16 − .216 = 1.944
Full page	2.48 − .248 = 2.232
Sum	5.176

This way of figuring out the results shows that the climax order, starting small and increasing the size, is decidedly less effective than the anticlimax order, an interesting difference between the art and the science of advertising.

It will be interesting to determine, on the basis of these figures, what we should expect as the result of our experiment, and then show the values actually obtained. In the experiment, however, no attention was paid to the order in which the different sizes of advertisements appeared, so, to make the two sets of figures at all comparable, it is necessary to find the maximum and minimum expectancy values and then observe whether the results obtained in the experiment lie between the two curves or a short distance to one side. In the following set of tables will be given the probable values of the different combinations of four presentations of the various sizes used in the experiment (pp. 458–9).

If these tables are based upon a correct theory, it is demonstrated that the order in which mixed sizes are presented will have a considerable effect upon the memory value of the advertisements of a firm. Since there is a close correlation between the memory value of mail order advertisements and their inquiry-producing power, it appears likewise that the actual effectiveness of a campaign will depend upon the order of presentation of the different sizes of advertisements which are used. From the tables it is possible to draw two quite definite conclusions.

1. The maximum efficiency is obtained when the largest advertisements appear first, followed by the smaller ones in descending order of size.

2. The minimum effectiveness is produced when the smallest advertisements are seen first, followed by the larger in ascending order of size.

These points ought to be of considerable value to the practical advertiser, likewise to the educator, for they demonstrate that the same amount of stimulation may produce vastly different results, one method showing a gain of from 10 per cent. to 20 per cent. in efficiency, the other a corresponding loss.

TABLE I

	1	2	3	4
3 full—1 half:				
FFFH	2.57	6.00	8.48	9.97
FFHF	2.57	6.00	7.73	9.64
FHFF	2.57	4.95	7.18	9.19
HFFF	1.62	4.71	7.06	9.13
3 full—1 quarter:				
FFFQ	2.57	6.00	8.48	9.48
FFQF	2.57	6.00	7.16	8.86
FQFF	2.57	4.17	6.15	8.06
QFFF	1.00	3.74	5.97	7.98
2 full—2 half:				
FFHH	2.57	6.00	7.73	9.15
FHFH	2.57	4.95	7.18	8.67
FHHF	2.57	4.95	6.60	8.51
HHFF	1.62	3.78	6.01	8.02
HFHF	1.62	4.71	6.44	8.35
HFFH	1.62	4.71	7.07	8.56
2 full—2 quarter:				
FFQQ	2.57	6.00	7.16	8.07
FQFQ	2.57	4.17	6.15	7.15
FQQF	2.57	4.17	5.24	6.94
QQFF	1.00	2.33	4.31	6.22
QFQF	1.00	3.74	4.90	6.60
QFFQ	1.00	3.74	5.97	6.97
2 full—1 half—1 quarter:				
FFHQ	2.57	6.00	7.73	8.64
FFQH	2.57	6.00	7.16	8.38
FHQF	2.57	4.95	6.02	7.72
FQHF	2.57	4.17	5.58	7.28
FHQF	2.57	4.95	7.18	8.18
FQFH	2.57	4.17	6.15	7.64
HFFQ	1.62	4.71	7.07	8.07
HFQF	1.62	4.71	5.87	7.57
HQFF	1.62	3.08	5.06	6.97
QHFF	1.00	2.94	5.17	7.08
QFHF	1.00	3.74	5.47	7.38
QFFH	1.00	3.74	5.97	7.46
1 full—2 half—1 quarter:				
FHHQ	2.57	4.95	6.60	7.51
FHQH	2.57	4.95	6.02	7.24
FQHH	2.57	4.17	5.58	6.86
HHFQ	1.62	3.78	6.01	7.01
HHQF	1.62	3.78	4.85	6.55
HFHQ	1.62	4.71	6.44	7.35
HQHF	1.62	3.08	4.49	6.40
HFQH	1.62	4.71	5.87	7.09
HQFH	1.62	3.08	5.06	6.55
QHHF	1.00	2.94	4.43	6.34
QHFH	1.00	2.94	5.17	6.66
QFHH	1.00	3.74	5.47	6.89

	1	2	3	4
1 full—1 half—2 quarter:				
FHQQ	2.57	4.95	6.02	6.89
FQHQ	2.57	4.17	5.58	6.49
FQQH	2.57	4.17	5.24	6.46
HFQQ	1.62	4.71	5.87	6.78
HQFQ	1.62	3.08	5.06	6.06
HQQF	1.62	3.08	4.15	5.85
QQHF	1.00	2.33	3.74	5.65
QQFH	1.00	2.33	4.31	5.80
QHFQ	1.00	2.94	5.17	6.17
QHQF	1.00	2.94	4.01	5.71
QFHQ	1.00	3.74	5.47	6.38
QFQH	1.00	3.74	4.90	6.12
1 full—3 half:				
FHHH	2.57	4.95	6.60	7.98
HFHH	1.62	4.71	6.44	7.86
HHFH	1.62	3.78	6.01	7.50
HHHF	1.62	3.78	5.35	7.26
1 full—3 quarter:				
FQQQ	2.57	4.17	5.24	6.11
QFQQ	1.00	3.74	4.90	5.81
QQFQ	1.00	2.33	4.31	5.31
QQQF	1.00	2.33	3.30	5.00
3 half—1 quarter:				
HHHQ	1.62	3.78	5.35	6.26
HHQH	1.62	3.78	4.85	6.07
HQHH	1.62	3.08	4.49	5.77
QHHH	1.00	2.94	4.43	5.74
2 half—2 quarter:				
HHQQ	1.62	3.78	4.85	5.72
HQHQ	1.62	3.08	4.49	5.40
HQQH	1.62	3.08	4.10	5.32
QQHH	1.00	2.33	3.74	5.02
QHQH	1.00	2.94	4.01	5.23
QHHQ	1.00	2.94	4.43	5.34
1 half—3 quarter:				
HQQQ	1.62	3.08	4.10	4.94
QHQQ	1.00	2.94	4.01	4.88
QQHQ	1.00	2.33	3.74	4.65
QQQH	1.00	2.33	3.30	4.52

We now return to a consideration of the experiment. It will be recalled that when it was started, there was no idea that the order of presentation would have any effect upon the outcome, so in the arrangement of the dummy, no attention was paid to this point. Consequently, in giving the results and showing the relation between the obtained and theoretical values, the maximum and minimum theoretical figures and the obtained ones will be shown, first in a table and then in a curve. The table follows:

Arrangement of Advertisements	Ratios of Av. Number of Mentions	A. D.[1]	Theoretical	
			Max.	Min.
3 full—1 half	9.55	2.50	9.97	9.13
3 full—1 quarter	8.57	8.00	9.48	7.98
2 full—2 half	7.81	8.75	9.15	8.02
2 full—2 quarter	7.16	5.00	8.07	6.97
2 full—1 half—1 quarter	8.41	1.75	8.64	7.07
1 full—2 half—1 quarter	5.03	5.25	7.51	6.34
1 full—1 half—2 quarter	5.52	6.50	6.89	5.65
1 full—3 half	7.57	3.00	7.98	7.26
1 full—3 quarter	6.26	2.50	6.11	5.00
3 half—1 quarter	6.51	6.00	6.26	5.74
2 half—2 quarter	5.03	4.00	5.72	5.34
1 half—3 quarter	4.12	1.75	4.94	4.52
1 full	2.12	1.60	2.57	2.57

In curve I. the same results are shown. When the values are plotted graphically, the very general similarity of the expected and the obtained results is clear. All but one of the points on the experimental curve lie close enough to the theoretical curve to be explained by the probable error of the experiment, which must be moderately high with only 35 subjects and 4 firms for each arrangement of advertisements. In the two curves, also, the directional change is the same in every place but one. While the results of the experiment do not absolutely prove the theoretical point, they render it extremely probable and open up an entirely new field for experimentation in the psychology of advertising.

An additional confirmation, while not completely conclusive, still points in the same direction. Shryer[2] gives, scattered through his book, a list of 9 insertions of advertisements followed by a second which was, judging from the

[1] It was pointed out by Dr. Froeberg that the A. D., in such an experiment, is practically useless. The A. D. may be worked out in either of two ways: First, by obtaining the average memory value for each advertisement and its A. D. By this method, the full page advertisements would have the smallest A. D., the half page next, and the quarter page advertisements the largest, for they are remembered by the fewest persons. Second, it is possible to group together all the firms advertising with 3 full and 1 half pages, in this case 4, obtaining the average and the A. D. This is the method which was employed, simply because some measure of variability is demanded in experimental work. But the A. D., to be of any value, presupposes material which is homogeneous, of equal difficulty. In these experiments, the advertisements are not of equal difficulty except by chance. Consequently, the A. D. is a product of two independent factors; the difficulty of the material, and the variability of the subjects.

[2] Shryer, W. A., 'Analytical Advertising.'

cost of space, 2.25 times the area of the first. Determining the probability in terms of his statistics for size and frequency, we find the expectancy value of the two insertions to be 2.575,

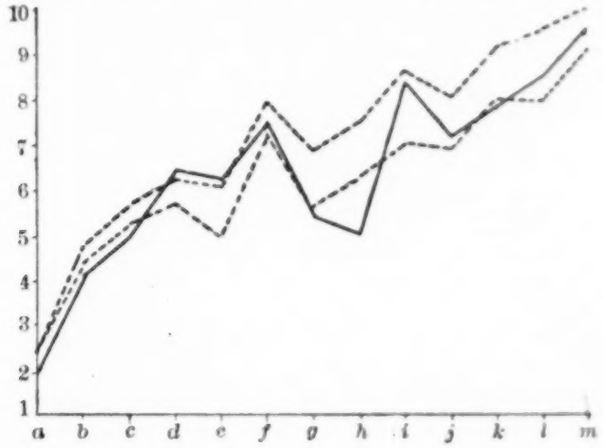

FIG. 1. Solid line: experimental results. Dotted line: maximum and minimum theoretical results.

a = 1 full,	*g* = 1 full—1 half—2 quarter,
b = 1 half—3 quarter,	*h* = 1 full—2 half—1 quarter,
c = 2 half—2 quarter,	*i* = 2 full—1 half—1 quarter,
d = 3 half—1 quarter,	*j* = 2 full—2 quarter,
e = 1 full—3 quarter,	*k* = 2 full—2 half,
f = 1 full—3 half,	*l* = 3 full—1 quarter,

m = 3 full—1 half.

and the actual value obtained from running the advertisements, 2.6000.

Still another bit of indirect confirmation of the theory is offered in curve II.

Fig. 2, *B*, shows the results obtained by using the values for size and frequency of insertion secured by certain mail order campaigns. Size was shown to have an efficiency of 1.00, 1.53 and 2.25 respectively for quarter, half and full page. Likewise, 1, 2, 3 and 4 insertions were discovered to give the following ratios: 1.00, 2.08, 2.78, 3.46. Using exactly the same method that was employed in the earlier parts of the paper, curve *B* was obtained. This curve is not guess work; it is based upon actual returns and a sufficient number of them to make it seem quite trustworthy. In giving results for frequency of insertion of advertisements in business, it is impossible to differentiate absolutely between duplication and variation, for both are probably effective. Fig. 2, *B*, then, would represent a mixture of the two.

Fig. 2, *A*, gives again the maximum theoretical results for variation, *C* for duplication worked out in terms of the tables on pages 458–9. Since curve *B* is a product of both of these factors, it ought to lie between *A* and *C*, and it does, almost equidistant from each all the way through, showing the same peculiarities, the same changes in direction, in fact, exhibiting the same tendencies throughout. The striking fact is the great degree of similarity in the forms of the various curves. The hypothetical ones, no matter whether for varia-

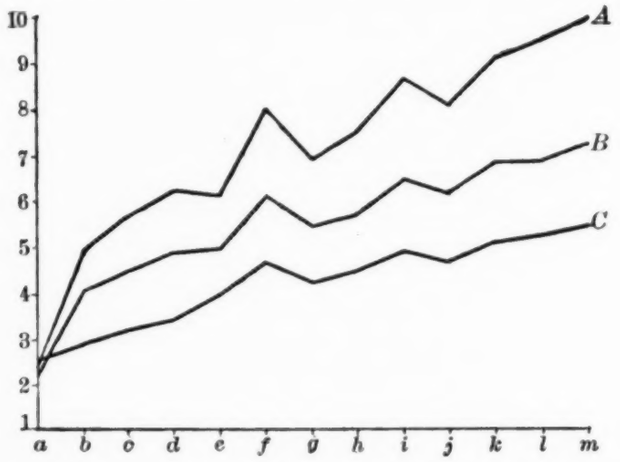

FIG. 2. The letters represent the same combinations as in Fig. 1.

tion or duplication, show the same general tendencies, though with differences of amount, as would be expected. The practical curve, *B*, is, to all intents and purposes, the mean between the two theoretical curves and the experimental curve is a very close approximation to *A*.

We may take it, then, that our main thesis is demonstrated in three ways, theoretically, practically and experimentally, viz., that using mixed sizes of advertisements may be made much more effective than using advertisements all of the same size. The second thesis seems almost equally probable, viz., that the order of distribution of the different sizes plays a large part in determining the efficiency of the total space employed.

Since there is such a close resemblance between the experimental results and the expectation curve, we shall use the latter as a basis for drawing further conclusions.

We do this because the expectation curve is based upon the results of a tremendous number of subjects as compared with the number used in our experiment.

If we take the combinations of sizes and show their relative effectiveness for different total areas according to the maximum theoretical curve, the following table is procured:

1 page	
4 quarters	4.13
1 half—2 quarters	4.10
1 full	2.57
1¼ pages	
1 half—3 quarter	4.94
2 half—1 quarter	4.85
1 full—1 quarter	4.17
1½ pages	
3 half	5.35
2 half—2 quarter	5.72
1 full—1 half	4.95
1 full—2 quarter	5.24
1¾ pages	
3 half—1 quarter	6.26
1 full—1 half—1 quarter	6.02
1 full—3 quarter	6.11
2 pages	
4 half	6.70
1 full—2 half	6.60
1 full—1 half—2 quarter	6.89
2 full	6.00
2¼ pages	
2 full—1 quarter	7.16
1 full—2 half—1 quarter	7.51
2½ pages	
2 full—2 quarter	8.07
2 full—1 half	7.73

There is no need to carry the table further, because all the other combinations are given in the preceding tables or curves. A brief inspection of the figures will show the most efficient combinations. The general conclusion is that when

variations are used, repetition is much more effective than size, and greater effectiveness will probably be obtained by using small, varied advertisements frequently inserted than by employing larger ones with less frequency. This is, of course, from the theoretical standpoint and takes no account of such considerations as the added cost of preparing new advertisements.

Where duplication is employed, provided anyone ever does use duplicated advertisements of different sizes, quite different results would be reached, for it will be remembered that with them, size is more effective than repetition. The exact amount of difference in any combination can be figured out by preparing tables like those given above.

In the practical situation, most advertising is a mixture of variation and duplication. From the standpoint of memory, what the effects of such a mixture will be is entirely problematical, but should be capable of experimental investigation with little difficulty. Certain practical results have already been given in Fig. 2, *B*, and it is probable that the memory curve would run closely parallel to it.

In conclusion, the results of psychological tests on advertisements, the results of the mail order campaigns which are accessible, these show that the effects of the objective factors follow rather simple root curves. The actual elevation differs somewhat. The results of certain other advertising campaigns, when plotted, follow definite power curves, the difference in the form of the curve resulting from differences in the kind of commodity.

These power and root curves are obtained by averaging or combining large numbers of records. It is not claimed that any one, or any dozen, insertions will lie along the curve, but in the long run, it will be approximated. This suggests a comparison with the tables which are used by life-insurance companies, which will not hold in any one case except by accident, but which, on the average, enable the life-insurance companies to make tremendous profits. It seems possible that advertising may in time become as scientific and as definite as insurance. But first, the advertisers must devise

some means of measuring the efficiency of their advertisements. Could they determine whether, for them, size and frequency followed definite power or root curves, they would have made a definite advance and could undoubtedly avoid considerable waste in conducting their campaigns. Even if they could not find root or power curves for the effectiveness of their advertisements, could they demonstrate the relative efficiency of mixed sizes, or mixed duplication and variation, they ought to be able to infer from their tables the type of curve which their advertising follows.

CHILDREN'S SENSE OF HARMONIES IN COLORS AND TONES

BY J. F. DASHIELL

Oberlin College

OBJECT

Many students have tested children's ability to perceive colors,[1] but only a few have studied their affective preferences in single colors,[2] and fewer still their preferences in color combinations.[3] So far as the writer is aware, no systematic study has been made of children's preferences in musical tone combinations or intervals. A preliminary object of the present study was first to get the writer's own results on the subjects here observed as regards single color judgments to check up with other investigations. But the main interest was in learning various points about *children's capacity to judge in terms of color combinations or harmonies;* and lastly to treat in a similar way if possible *their judgments of tone intervals or harmonies.*

MATERIALS AND METHODS

212 kindergarten children were used as the subjects, taken from nine Minneapolis kindergartens, both public and private, and including children of all classes. The experiments were given with the help of teachers-in-training. For purposes of comparison, the same observations were made upon 126 sophomores in the writer's classes in general psychology at the University of Minnesota.

[1] Baldwin, Binet, Bullough, Garbini, Lobsien, McDougall, Marsden, Monroe, Myers, Nagel, Preyer, Shinn, Wolfe, Ziehen.

[2] Engelsperger and Ziegler, 'Weitere Beiträge z. Kenntnis d. phys. u. psych. Natur d. sechsjährigen in d. Schule eintretenden Kindes,' *Experimentelle Pädagogik*, 1906, **2**, 49–95. Holden and Bosse, *Archiv. of Ophthal.*, 1900, **29**, number 3. Winch, W. H., Color Preferences of School Children, *Brit. J. Psychol.*, 1909, **3**, 42–65.

[3] Dobbie, W. J., Experiments with School Children on Color Combinations, *Univ. of Toronto Studies*, 1900, **1**, 253–267. Wundt, W., Die Empfindung d. Lichts u. d. Farben, *Phil. Studien*, 1888, **4**, 311–389.

In the investigation of single color preferences the six standard kindergarten colors (Milton Bradley Co.) were used —red, orange, yellow, green, blue, violet, corresponding, in Ridgway's classification, to colors 2i, 8–, 21–, 34–, 51–, 61h, respectively.[1] Rectangles of 13 mm. by 45 mm. were cut from paper of the six colors and pasted in the centers of neutral gray (Milton Bradley Co.) cardboards, 80 mm. by 90 mm. A modified order-of-merit method was adopted for the test, the child picking out the card bearing the color liked best from the whole number on the table, the experimenter removing it, the child picking the best of those remaining, etc. In the study of preferences in color combinations the same kind of cards was used, with same colors mounted in pairs, there being six pairs used—red-green, red-blue, orange-yellow, orange-green, yellow-violet, blue-violet. The order-of-merit method was again adapted to the situation.

In the experiments on tone harmony preferences, the experimenter, seated at the piano, sounded five musical intervals one at a time, asking before and after each, "Do you like this sound?" or "Do you think this a pretty sound?" Repetitions were allowed. The child's judgments in 'Yes' or 'No' were recorded for each interval. The intervals used were:

Major third,	c^1–e^1,	264–330	vibrations;
" seventh,	c^1–b^1,	264–495	" ;
" fifth,	c^1–g^1,	264–396	" ;
Octave,	c^1–c^2,	264–528	" ;
Minor second,	c^1–$d^{\flat 1}$,	264–281.6	" .

RESULTS

I

The Preliminary Test for Single Color Preferences.—In the tables is shown first the total number of times each color was given first choice, second choice, etc. Then, by multiplying the number of times a given color is placed in first

[1] R. Ridgway, "*Color Standards and Color Nomenclature.*" published by author, Washington, 1912.

rank by the figure 1, in second rank by the figure 2, etc., and adding the products, one obtains a figure representing inversely that color's value as resulting from all its placings. This is given as the total value (T. V.). On the basis of the total values, the six different colors can then be given their final order of preference or rank (R.).

TABLE I

Kindergarten Girls and Boys (212)								
	1	2	3	4	5	6	T. V.	R.
Red	45	34	29	31	37	36	725	2
Orange	28	36	37	35	40	36	767	6
Yellow	36	31	41	29	36	39	751	3
Green	29	33	38	42	37	33	760	5
Blue	40	42	34	42	28	26	690	1
Violet	34	36	33	33	34	42	759	4

The indefiniteness and small reliability of such measurements is at once evident, as in the difference between value 690 for blue (ranked first) and value 767 for orange (ranked sixth). One result of interest aside from the very slight differences given in the table is that the table for girls only and that for boys only shows agreement in giving first place to blue and fourth place to yellow. The difference of color given second place is striking.

TABLE II

Kindergarten Girls (107)			Kindergarten Boys (105)		
	T. V.	R.		T. V.	R.
Red	381	5	Red	344	2
Orange	416	6	Orange	351	3
Yellow	374	4	Yellow	377	4
Green	367	3	Green	393	5
Blue	348	1	Blue	342	1
Violet	361	2	Violet	398	6

The tables in a general way are in some agreement with the results of earlier studies.[1] The placing of red by the

[1] Winch for girls in (English) standards II. to VII. gives the following order of preference: blue, red, white, green, yellow, black; for boys: blue, red, yellow, green, white, black. Holden and Bosse find red preferred by infants to 3 years of age, with blue predominating more and more thereafter. Engelsperger and Ziegler find 6-year-olds giving decided preference to purple, dark blue, and violet.

girls is the only surprising fact. This may be due in part to the particular shade of red used. Later it will be observed that the college girls restore red to its expected place.

The correct interpretation of the close grouping of the total values for the respective colors would be important. This may be an indication that the order-of-merit method is not adequate for studying preferences of children. On the other hand, it more probably indicates that preferences in young childhood, however decided or undecided at the time, are comparatively irregular and probably inconstant, and are to be understood best in connection with the instability in all phases of childhood. It is further evidence for the lack of mental and nervous organization that is to be recognized as fundamental here.

When we compare the results above with the results of similar observations upon college sophomores interesting points come out.

TABLE III

Sophomore Girls and Boys (126)

	1	2	3	4	5	6	T. V.	R.
Red	38	26	17	19	17	9	356	2
Orange	7	12	17	21	27	42	553	6
Yellow	9	16	12	31	31	27	508	5
Green	20	19	31	31	16	9	409	3
Blue	36	32	24	13	11	10	339	1
Violet	16	22	24	11	24	29	470	4

A very obvious difference is that here there is a more clear-cut order of preference, the total values of the six colors showing larger differences between them. Thus, between the 339 for blue (ranked first) and the 553 for orange (ranked sixth) we find a more substantial difference than between 690 and 767 above. Is this a matter of self-organization through experience, making one's maturer judgments proceed from a firmer physiological and psychological basis, or is it a matter of social control in the form of innumerable social agencies that tend to make individuals uniform? Put differently, the question is: is this greater uniformity indicative of a similarity of structure and innate tendency that shows itself with time, or is it indicative merely of the similarity of social forces that help determine development?

Aside from this difference of reliability, it is interesting to observe a fairly close correlation between the order of preference given by the sophomores and that by the child. With the exception of yellow and green, which are reversed, the orders are exactly the same: Blue, red, green or yellow, violet, yellow or green, orange. This might add a little strength to the first type of theory referred to.

Sex differences among the sophomores were indicated: the difference between the choice of red or blue for first and second place, and the difference between orange and violet for fourth and sixth.

TABLE IV

Sophomore Girls (92)			Sophomore Boys (34)		
	T. V.	R.		T. V.	R.
Red	255	1[1]	Red	101	2
Orange	419	6	Orange	134	4
Yellow	361	5	Yellow	147	5
Green	304	3	Green	105	3
Blue	271	2	Blue	68	1
Violet	321	4	Violet	149	6[2]

Comparing Tables IV. and II., the sex differences appear in general similar at the two ages, with the striking exception of the girls' low ranking of red in kindergarten and high ranking in college.

II

In the test for color combinations the kindergarten children showed decidedly little definiteness of choice.

TABLE V

Kindergarten Girls and Boys (210)								
	1	2	3	4	5	6	T. V.	R.
Red-green	47	29	35	34	33	32	703	1
Red-blue	38	35	30	44	37	26	715	2
Orange-yellow	42	35	35	26	31	41	722	3½
Orange-green	18	33	42	42	43	32	785	6
Yellow-violet	31	46	32	35	33	33	722	3½
Blue-violet	34	32	36	29	33	46	763	5

[1] Compare with Winch's observation that 41 women teachers ranked red low: blue, green, white, red, yellow, black.

[2] Compare Jastrow's report (Popular Æsthetics of Color, *Pop. Sci. Mo.*, 1897, 50, 361–368) that adult men rank blue decidedly first with red second and that adult women rank red first with blue second.

Differences between boys and girls show in the placing of the first three cobminations.

TABLE VI

Kindergarten Girls (106)			Kindergarten Boys (104)		
	T. V.	R.		T. V.	R.
Red-green	368	4	Red-green	335	1
Red-blue	360	1	Red-blue	355	2
Orange-yellow	363	2	Orange-yellow	359	4
Orange-green	394	6	Orange-green	391	6
Yellow-violet	366	3	Yellow-violet	356	3
Blue-violet	375	5	Blue-violet	388	5

The extremely close grouping of the children's judgments of the color combinations is the striking fact of this set of observations. And as one casts his eye over the detailed results he is struck even more by the same thing. The children were grouped according to their nine different kindergartens and according to sex; and the eighteen resulting group summaries bear almost no resemblance to each other, each color combination having been given each of the six ranks at least once. That there is practically no agreement among young children as to their reactions to combinations of colors would seem, then, to be indicated. Now, again, whether this lack of a common basis is due merely to undeveloped, unorganized character, or is due to the limited amount and varying nature of the social control by suggestion, would make an interesting problem. That in *some way* it is due to immaturity, is, however, established by a comparison with the results from the college students.

TABLE VII

Sophomore Girls and Boys (125)								
	1	2	3	4	5	6	T. V.	R.
Red-green	10	17	27	25	21	25	480	5
Red-blue	30	26	22	20	15	12	375	1
Orange-yellow	18	16	27	28	27	9	442	4
Orange-green	8	15	16	16	31	39	539	6
Yellow-violet	41	19	15	14	18	18	378	2
Blue-violet	18	32	18	22	13	22	421	3

That the preferences are again more uniform and more

definite than with children is exhibited, *e. g.*, by the difference between T. V. 375 for red-blue (ranked first) and 539 for orange-green (ranked sixth).

In the matter of similarities between actual choices made by children and college students, perhaps the only correlations of importance are in ranking the orange-green lowest and the red-blue at or near the top.

Sex differences among the sophomores are interesting.

TABLE VIII

Sophomore Girls (91)			Sophomore Boys (34)		
	T. V.	R.		T. V.	R.
Red-green	365	5	Red-green	115	3
Red-blue	295	3	Red-blue	80	1
Orange-yellow	328	4	Orange-yellow	114	2
Orange-green	421	6	Orange-green	118	4
Yellow-violet	227	1	Yellow-violet	151	6
Blue-violet	285	2	Blue-violet	136	5[1]

The girls show a preference for the yellow-violet combination with blue-violet and red-blue rated good; while the boys show a strikingly consistent preference for the red-blue, rating the other two just mentioned low. And this difference between the sexes is clear in the summaries for each of the three sections into which the students were divided.

That the sex differences here are not similar to those at the kindergarten age (Table VI.) would logically follow from the general dissimilarity, noted above, in ranking by boys and girls together at the two ages (Tables V. and VII.).[2]

III

In the experiment on tone-intervals the favorable or unfavorable reaction as expressed by the child for each

[1] Jastrow got a somewhat different result with his adults. Using a greater range of color combinations he found that women place light red-light green first, red-green second, red-light green third; men place red-blue first, blue-violet second.

[2] The studies of Dobbie and of Wundt on children's preferences in color combinations were concerned with a somewhat different problem. Dobbie shows that "the combinations with minimum difference (spectral order) [between successive colors] has a decided preference." This confirms Wundt's finding that his two children found most pleasing those combinations which showed spectral proximity and order.

interval sounded was recorded by a positive (+) or negative (−) sign respectively.

TABLE IX

	Kindergarten Girls (99)		Kindergarten Boys (99)		Both (198)	
	+	−	+	−	+	−
c—e	97	2	96	3	193	5
c—b	69	30	63	36	132	66
c—g	91	8	93	6	184	14
c—c	89	10	87	12	176	22
c—d^b	53	46	58	41	111	87

Compare this with the table for the sophomores.

TABLE X

	Sophomore Girls (63)		Sophomore Boys (26)		Both (89)	
	+	−	+	−	+	−
c—e	59	4	25	1	84	5
c—b	—	63	—	26	—	89
c—g	60	3	21	5	81	8
c—c	60	3	25	1	85	4
c—d^b	1	62	—	26	1	88

Again, the striking difference between the two ages of subjects is one of definiteness and reliability of results. As was to be expected from the simplicity of the intervals the college students reacted almost unanimously. The intervals are well recognized as pleasant or unpleasant by adults in general. But the children's reactions, with such great variability in their likes and dislikes for such intervals, may be a bit unexpected.

What is to be made of this? (*A*) Shall we infer that the comparatively unorganized character of child-nature, the absence as yet of æsthetic sense, is the complete explanation: that all that is necessary to get results as in Table X. from subjects of Table IX. is growth and natural mental development? Such an interpretation would seem to imply that by their very physical nature and the relationship of that to the physiological nature of the adult human subjects, intervals c–b and c–d^b are and must be such as to arouse unpleasant

reactions, c–e, c–g, and c–c, pleasant reactions. This more objective and absolute interpretation must take into consideration the character of the component physical vibrations, and whatever is to be known about the physiological effects of differing combinations of vibration frequencies. (*B*) Or, shall we insist that something of an æsthetic sense is already present, but that it is modifiable, plastic; and that between the kindergarten and the college ages, come the multiplicity of social suggestions that tend to make all individuals of a society more or less alike in their standards of what is and what is not likable? To put this point more concretely: it is conceivable, at least, that if put in a social community in which the intervals c–b and c–$d^{\flat}$ were handled as pleasant intervals and the musical forms of such a community based upon the pleasantness of these and the unpleasantness of the c–e, c–g, and c–c intervals, the child would develop a set of preferences totally different from those shown by the college students above.

And, of course, the (*A*) and (*B*) interpretations are quite relevant to the matter of differences in color preferences as set forth in the earlier paragraphs of this paper.

Summary

In the experiments on single colors and on tone intervals but not on color combinations there is a rough qualitative agreement between the order of judgments made by the two ages.

With single colors the two sexes at both ages show a similarity in preferring certain colors but a real difference in the order in which they are preferred. With the exception of red—which seems to involve an unknown disturbing factor—age makes little difference. With color combinations sex comparisons are valueless with the children, but with the sophomores they indicate preference for red-blue by boys and yellow-violet by girls. With tone intervals there are no sex differences brought out.

The outstanding results of all three sets of experiments is the clear demonstration of a decidedly greater variability

and unreliability (for group purposes, at least) of æsthetic judgments by children of kindergarten age as compared with those by later adolescents and adults. What this may be due to has been suggested as a fundamental problem.

Pedagogical applications would seem to follow. In general, it would appear that kindergarten methods which are now necessarily based more or less on *a priori* foundations and on general observations might benefit from critical studies of how the child does actually react to æsthetic (and other) stimuli.

THE RELATION BETWEEN LEARNING AND RETENTION AND AMOUNT TO BE LEARNED[1]

BY V. A. C. HENMON

University of Wisconsin

In the experimental study of memory the first problem taken up was the relation between the length of series of nonsense-syllables and the number of repetitions required for learning them. Does the difficulty of learning increase proportionally with an increase in the number of syllables, more slowly or more rapidly? Ebbinghaus[2] in the widely quoted

Ebbinghaus (Grundzüge)	Syllables	6	12	16	36
	Repetitions	1	14–16	30	55
James	Syllables	7	12	24	26
	Repetitions	1	16	44	55
Meumann	Syllables	7	12	16	24–36
	Repetitions	1	16.6	30	44–65
Claparéde	Syllables		12	24	36
	Repetitions		16.5	44	55
Myers	Syllables	7	12–16	24	36
	Repetitions	1	16.6–30	44	55

results found that while 7 syllables could be recalled with one presentation, 12 syllables required 16.6 repetitions, 16 syllables 30 repetitions, 24 syllables 44 repetitions, and 36 syllables 55 repetitions. In other words, the number of repetitions increases at first with very great rapidity and then less rapidly, the increase in repetitions being relatively much greater than the increase in the number of syllables. Binet and Henri[3] found a similar tendency in the learning of num-

[1] Paper read before the American Psychological Association at the Washington meeting, 1911. Publication has been delayed in the expectation that further work could be done on the problem. In view of the full report of the paper by Lyon, D. O., *J. of Educ. Psychol.*, 1914, 5, pp. 7–9, and the review by Pieron, H., *Année Psychol.*, Vol. XIX., 1913, pp. 119–120, it seems best to publish the paper as read.

[2] Ebbinghaus, H., 'Über das Gedächtnis,' p. 64. 'Grundzüge der Psychologie, pp. 650–651. The figures for these experiments have been often and variously reported:

[3] Binet, A., & Henri, V., 'La Memoire des Mots,' *Année Psychol.*, 1895, 1, p. 12.

bers. The increase in time for learning with the increase in length of series was as follows:

10 numbers	.....	17 secs.
15 "	.....	75 "
20 "	.....	135 "
25 "	.....	180 "
30 "	.....	260 "
50 "	.....	420 "
100 "	.....	1,500 "
200 "	.....	4,520 "

The explanation of this result is sought by Ebbinghaus in the narrowness of the span of consciousness and in retroactive inhibition. The narrowness of the span of consciousness accounts for the increase in repetitions with twelve syllables over seven. Retroactive inhibition accounts for the increase with the longer series. This retroactive inhibition is exerted by each individual association not only upon those immediately preceding it, but to some extent on all those more remote as well. The effect, therefore, is greater the greater the number of inhibiting and inhibited members. Offner[1] suggests as a partial explanation that the greater the number of members in a series the less attention given to each member, apparently assuming that the amount of attention to any series is a constant quantity and that in a long series it is merely more widely distributed; hence, the greater number of repetitions for learning. He finds a second factor in the influence of forgetting which begins immediately and has progressed so much further in a long series than in a short one that it requires numerous repetitions to bring it back to the level of a short series. Myers[2] attributes the result to fatigue and the lesser concentration of attention with the longer series.

No systematic inquiry has been made to test the validity of Ebbinghaus's law and his data are quoted in all the books. This seems curious in view of the fact that Ebbinghaus expressly states that his results are tentative and that he

[1] Offner, M., 'Das Gedächtnis,' Berlin, 1909, p. 54–55.

[2] Myers, C. S., 'Text-Book of Experimental Psychology,' London, 1909, pp. 157–158.

does not regard them as very reliable. More recent investigations, however, so far as they have gone, have failed to confirm them.

Radossawljewitch[1] noted in his experiments on the rate of forgetting that the change from 8 to 12 or to 16 syllable series did not demand a great increase in the number of repetitions, in fact, frequently 16 syllables were memorized with fewer repetitions than 8 or 12 syllables. To test the matter more fully Meumann learned 5 series each of 8 and 12 syllables and 2 series each of 16, 18, 24 and 36 syllables with the following results:

No. of Syllables	Repetitions	Differences
8	5.2	
12	10.4	5.2
16	17.0	6.6
18	21.5	4.5
24	30.0	8.5
36	32.5	2.5

These results are the reverse of those obtained by Ebbinghaus. There is a relative decrease in repetition with an increase in the number of syllables. Meumann[2] holds that this is what might be expected. An increase in amount of work to be done, if it is not too great, makes little difference when once the initial disinclination or inertia is overcome, when adaptation of attention is secured, when the associative processes have been aroused and a general adjustment to the work is once attained. All of these formal conditions of learning should be effective for a series no matter what their length within the limits of fatigue. Hence it is reasonable to expect a relative decrease in energy required for learning with an increase in amount to be learned.

The relation between amount to be learned and retention as distinguished from learning is not clearly made out. Ebbinghaus found that the longer the series the more firmly it is impressed; hence after 24 hours both the absolute and

[1] Radossawljewitch, P., 'Das Behalten und Vergessen bei Kindern und Erwachsenen,' Leipzig, 1906.

[2] Meumann, E., 'Über Ökonomie und Technik des Gedächtnisses,' Leipzig, 1908.

relative economy is greatest with the longest series. The economy after twenty-four hours was as follows:

No. of Syllables	Per Cent. Saved.
12	33
24	49
36	58

Binet and Henri[1] and Pentchew[2] confirmed this result. Radossawljewitch[3] also notes that the shorter series were more quickly forgotten than the longer series except for the intervals of 5 minutes and 20 minutes.

Müller and Pilzecker and Reuther, however, claim that while the absolute amount retained increases with the increase in the length of the series, the relative amount decreases.

Ebbinghaus' results suggest that economy in mental work would result from a subdivision of it into small amounts. This flies in the face of common experience and the results of the experiments on learning by wholes and by parts. From Meumann and Radossawljewitch's results one would conclude that it is more economical to take large amounts at a time, the larger the amounts, the better within the limits of fatigue and available psychophysical energy.

Ebbinghaus' experiments are on one subject only. Meumann's results are also on one subject and the number of cases small, the figures for the longer series being averages from two series. Hence, it seemed worth while to continue the investigation on a number of individuals with nonsense-syllables and also with sense material, poetry and prose.

Experiments and Results

1. *Nonsense Syllables.*—The first experiment was designed to test Ebbinghaus and Meumann's results in the learning of nonsense-syllables. Eight series each of 10, 12, 14, 16, 18, 20, 24 and 30 syllables were prepared for exposure with the Ranschburg apparatus. These were presented to the subjects

[1] *Op. cit.*

[2] Pentchew, C., 'Untersuchungen zur Ökonomie und Technik des Lernens,' *Arch. für d. ges. Psychol.*, I, 1903.

[3] *Op. cit.*

at a rate which permitted the exposure of each syllable for 1¼ seconds. Two subjects took part in the experiments. Ordinarily, the subject worked for one hour at the same time a day and relearned after twenty-four hours.

The results of the experiment, for Subject *H*, appear in Table I. It gives the average number of repetitions required for learning, based on 8 experiments and for relearning based on 4–6 cases, the mean variations, and the per cent. of repetitions saved after twenty-four hours.

TABLE I.

	No. of Syllables							
	10	12	14	16	18	20	24	30
Repetitions for learning	6.9	8.1	8.3	8.6	10.6	13.6	12.9	20.0
M. V.	.9	1.5	1.1	1.6	1.6	2.4	1.5	3.3
Repetitions for relearning	3.8	4.5	4.8	4.4	5.5	6.5	5.7	9.5
M. V.	.4	.5	.4	.6	.5	.8	.7	.5
Per cent. saved after twenty-four hours	45	44	42	49	48	52	56	53

Table II. gives the results for Subject *C* for learning, the averages being from 3 experiments.

TABLE II.

	No. of Syllables							
	10	12	14	16	18	20	24	30
Repetitions for learning	12.7	14.0	15.3	15.0	16.3	19.0	16.0	25.7
M. V.	1.1	1.3	1.6	0.0	1.8	1.3	2.7	3.6

Fig. 1 indicates graphically the curve of increase in repetitions and for purposes of comparison the curves obtained by Ebbinghaus (E) and Meumann (M) are included.

The results of the experiments diverge widely from those of Ebbinghaus and Binet and are in fairly close accord with those of Meumann and Radossawljewitch. There is a relative decrease in the number of repetitions as the number of syllables increases. Particularly noteworthy is the fact that the number of repetitions for the series from 10 to 18 is practically constant. The results are even more striking than those of Meumann in showing the relative economy with

the longer series. It seems clear that Ebbinghaus' results do not represent a general tendency.

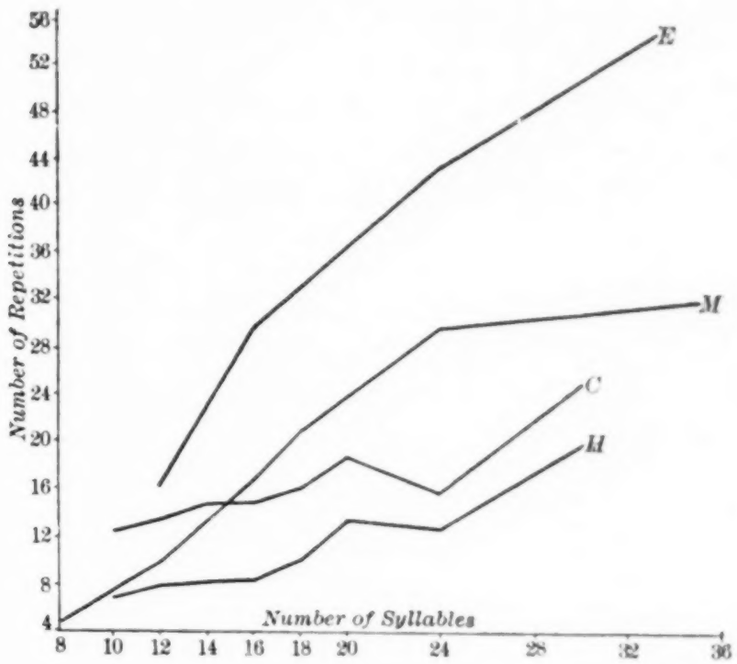

FIG. 1.

As is to be expected the saving in repetitions after twenty-four hours will not show the marked differences which Ebbinghaus obtained. Ebbinghaus found a saving of 33 per cent. in repetitions with 12 syllables after twenty-four hours, of 49 per cent. with 24 syllables, of 58 per cent. with 36 syllables. In other words, the force of association increases progressively with the increase in the number of syllables. In the results with Subject *H*, the saving after twenty-four hours is also slightly greater with the longer series. Both from the point of view of repetitions required for learning and for relearning after twenty-four hours the economy is greatest with the longer series.

Weber working in Meumann's laboratory reports that the law holds for subjects untrained in learning syllables. A repetition of the experiments on six untrained subjects with 10, 12, 14, 16, 18, 20 syllables fails to confirm this result. The data appear in Table III. (The figures in parentheses indicate the number of experiments.)

TABLE III.

	10	12	14	16	18	20
A (3)	17.7	28.7	22.7	22.7	22.3	38.3
B (2)	14.0	20.0	19.0	23.5		
C (2)	11.0	18.0	23.5	25.0		
D (1)	20.0	25.0	27.0	31.0		
E (2)	9.0	7.5	9.5	16.0		
F (1)	9.0	10.0	11.0	26.0		
Av.	13.4	18.2	18.8	24.0	22.3	38.3

None of these subjects had ever learned nonsense-syllables. To eliminate the influence of practice the order of the series varied with each subject, which accounts for the irregularities in the results, *e. g.*, *A* learned 12 syllables first, *F*, 16 syllables, etc.

2. *Poetry.*—The second set of experiments consisted in memorizing by the 'whole' method 1, 2, 3, 4 and 5 stanzas of 'In Memoriam,' a record being kept of the number of repetitions required for learning and for relearning after twenty-four hours. The results on three subjects, giving the repetitions for learning (L.), for relearning after twenty-four hours (R.), the per cent. saved and the mean variations appear in Table IV. The figures in parentheses give the number of experiments with each group of stanzas.

TABLE IV.

Stanzas	H (10)					D (10)					P. (5)					Average		
	L.	M. V.	R.	M. V.	Economy, Per Cent.	L.	M. V.	R.	M. V.	Economy, Per Cent.	L.	M. V.	R.	M. V.	Economy, Per Cent.	L.	R.	Economy, Per Cent.
1	3.3	.6	1.3	.4	61	3.5	.7	1.1	.4	69	3.8	.6	2.0	.4	47	3.5	1.5	57
2	5.3	.7	2.2	.3	59	6.1	.9	1.8	.5	70	7.4	1.5	3.0	.4	59	6.3	2.3	63
3	6.7	.9	2.3	.6	66	8.9	1.3	1.4	.9	84	10.2	1.8	3.8	.8	63	8.6	2.5	71
4	7.5	1.1	2.5	.5	67	10.7	1.9	1.9	.5	82	11.8	1.8	3.8	.6	68	10.0	2.7	73
5	8.6	1.2	2.6	.6	69	14.0	1.8	2.3	1.6	80	14.0	2.8	4.6	.7	67	12.2	3.2	74

The increase in the number of repetitions with the increase in amount is relatively less than the increase in the number of lines or stanzas. If the increase was proportional to the amount the number of repetitions would be 3.5, 7.0, 10.5,

14.0 and 17.5 instead of which the series is 3.5, 6.3, 8.6, 10.0, and 12.2. There is thus a relative economy with the larger amounts. The economy in relearning after twenty-four hours is greater with the larger amounts and is relatively greater with poetry than with nonsense syllables.

'In Memoriam' is rather difficult material for memorizing and the rate of increase is relatively rapid. One practiced subject with 1, 2, 3, 4, stanzas of 'Childe Harold' gave a less rapid increase.

Stanzas	L.	R.	Per Cent. Saved
1	8.3 Rep.	4.8 Rep.	40
2	9.3 "	5.5 "	40
3	10.0 "	5.4 "	46
4	9.1 "	4.8 "	47

A test of nine unpracticed subjects gave the following results:

No. of stanzas	1	2	3	4	5
Repetitions	4.6	9.7	11.8	15.9	15.3

3. *Prose.*—For the experiments on learning prose, 100-word, 200-word, and 300-word passages were selected from the essays of Huxley and Matthew Arnold. 60 selections from Huxley were made, and 60 from Arnold, 20 passages of each length. One practiced subject learned 54 of these selections, 18 from each group, and recorded the number of repetitions required for learning. Eight of each of the selections were relearned after twenty-four hours and 10 of each group were relearned after seven months. The rate of reading in learning was made as uniform as possible. The results are indicated in Table V.

TABLE V.

	100 Words	200 Words	300 Words
Learning Huxley passages	7.4 Rep.	7.9 Rep.	7.4 Rep.
Relearning after twenty-four hours	4.0	3.0	2.8
Per Cent. saved	46%	62%	62%
Learning Arnold passages	5.5 Rep.	6.7 Rep.	6.7 Rep.
Relearning after seven months	4.3	5.0	4.3
Per Cent. saved	21%	25%	35%

There is an approximate constancy in the number of repetitions for the passages irrespective of their length. The averages for the 18 passages of each length are:

Repetitions.....	100 words	200 words	300 words
	6.1	7.3	7.0

It may be remarked that a double fatigue order in learning was observed so as to eliminate inequalities in practice effect.

Experiments which have been made on the distribution of repetitions and on learning by wholes and by parts would lead one to infer that Ebbinghaus's law could not well represent a general tendency. My results show the inference to be correct. Retroactive inhibition, forgetting, the distribution of attention and fatigue which have been invoked to explain Ebbinghaus' results are more than compensated for by perseveration, by the time permitted for the setting of associations, and by the greater effort of attention which the longer series calls forth.

PSYCHOLOGICAL REVIEW PUBLICATIONS

Original contributions and discussions intended for the Psychological Review should be addressed to

Professor Howard C. Warren, Editor *Psychological Review*,
Princeton, New Jersey.

Original contributions and discussions intended for the Journal of Experimental Psychology should be addressed to

Professor John B. Watson, Editor *Journal of Experimental Psychology*,
The Johns Hopkins University, Baltimore, Maryland.

Contributions intended for the Psychological Monographs should be addressed to

Professor James R. Angell, Editor *Psychological Monographs*,
University of Chicago, Chicago, Illinois.

Reviews of books and articles intended for the Psychological Bulletin should be addressed to

Professor E. C. Tolman, Editor *Psychological Bulletin*,
University of California, Berkeley, California.

Titles and reprints intended for the Psychological Index should be addressed to

Professor Madison Bentley, Editor *Psychological Index*,
University of Cornell, Ithaca, New York.

All business communications should be addressed to

Psychological Review Company,
Princeton, New Jersey.

PSYCHOLOGICAL MONOGRAPHS

VOL. I

Aesthetic Experience: Its Nature and Function in Evolution. By Vernon Lee. Pp. xvi + 160. $2.00. 2. **The Philosophy of John Grote.** Edited by Alexander Macfarlane. Pp. viii + 131. $1.50. 3. **Emotion and Feeling: An Introduction to the Psychology of Dynamics.** James H. [illegible]. Pp. [illegible]. $2.00.

LIBRARY OF GENETIC SCIENCE AND PSYCHOLOGY

Vol. 1. **The Mind: Its Life.** A Study in Genetic Psychology. By Edwin A. Kirkpatrick, Professor of Psychology in Cornell University. Pp. [illegible]. $1.50 net. Postage 10 cents.

Vol. 2. **Heredity and Eugenics.** By Edwin G. Conklin, Professor of Biology in Princeton University. Pp. [illegible]. $1.50 net. Postage 10 cents.

Published by

PSYCHOLOGICAL REVIEW COMPANY

PRINCETON, N. J.